DOCUMENT RESUME

ED 085 406 TM 003 347

AUTHOR	Scanlon, James
TITLE	Intellectual Development of Youths As Measured by a Short Form of the Wechsler Intelligence Scale: United States.
INSTITUTION	National Center for Health Statistics (DHEW), Rockville, Md.
REPORT NO	DHEW-HRA-74-1610
PUB DATE	Sep 73
NOTE	50p.; Vital and Health Statistics, Series 11, n128
AVAILABLE FROM	Superintendent of Documents, U. S. Government Printing Office, Washington, DC 20402 ($0.80)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Age Differences; Analysis of Variance; *Intellectual Development; *Intelligence Differences; *Intelligence Tests; Sex Differences; Tables (Data); *Test Results; *Youth
IDENTIFIERS	*Wechsler Intelligence Scale for Children; WISC

ABSTRACT

This report summarizes the results of the administration of the vocabulary and Block Design subtests of the WISC to a national probability sample of noninstitutionalized youths 12 through 17 years of age. Information is presented on the derivation of percentile equivalents, normalized scale scores, and a short-form estimate of Face Scale Scores. An analysis of variations in raw scores, scales scores and estimates of full scores associated with age, sex, and grade placement is provided. (Author)

ED 085406

TM 003 347

VITAL and HEALTH STATISTICS

DATA FROM THE NATIONAL HEALTH SURVEY

Series 11-Number 128

NCHS NATIONAL CENTER FOR HEALTH STATISTICS

Intellectual Development of Youths

As Measured by a Short Form of the Wechsler Intelligence Scale

United States

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Health Resources Administration

Vital and Health Statistics-Series 11-No. 128

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 - Price 80 cents

ED 085406

Data from the
NATIONAL HEALTH SURVEY

Series 11
Number 128

Intellectual Development of Youths

As Measured by a Short Form of the Wechsler Intelligence Scale

United States

A discussion of results from the Vocabulary and Block Design subtests of the WISC administered to a national probability sample of noninstitutionalized youths 12 through 17 years of age. Description of the derivation of percentile equivalents, normalized scaled scores, and a short-form estimate of Full Scale Scores. Analysis of variations in raw scores, scaled scores, and estimates of Full Scale Scores associated with age, sex, and grade placement.

DHEW Publication No. (HRA) 74-1610

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service

Health Resources Administration
National Center for Health Statistics

Rockville, Md. September 1973

NATIONAL CENTER FOR HEALTH STATISTICS

EDWARD B. PERRIN, Ph.D., *Acting Director*
PHILIP S. LAWRENCE, Sc.D., *Deputy Director*
GAIL F. FISHER, *Assistant Director for Health Statistics Development*
WALT R. SIMMONS, M.A., *Assistant Director for Research and Scientific Development*
JOHN J. HANLON, M.D., *Medical Advisor*
JAMES E. KELLY, D.D.S., *Dental Advisor*
EDWARD E. MINTY, *Executive Officer*
ALICE HAYWOOD, *Information Officer*

DIVISION OF HEALTH EXAMINATION STATISTICS

ARTHUR J. McDOWELL, *Director*
GARRIE J. LOSEE, *Deputy Director*
PETER V. V. HAMILL, M.D., *Medical Advisor, Children and Youth Program*
HENRY W. MILLER, *Chief, Operations and Quality Control Branch*
LINCOLN I. OLIVER, *Chief, Psychological Statistics Branch*
HAROLD J. DUPUY, Ph.D., *Psychological Advisor*

COOPERATION OF THE BUREAU OF THE CENSUS

In accordance with specifications established by the National Health Survey, the Bureau of the Census, under a contractual agreement, participated in the design and selection of the sample, and carried out the first stage of the field interviewing and certain parts of the statistical processing.

Vital and Health Statistics-Series 11-No. 128
DHEW-Publication No. (HRA) 74-1610
Library of Congress Catalog Card Number 72-600141

CONTENTS

	Page
Introduction	1
General	1
Source of the Data	1
Psychological Test Battery	2
The Wechsler Intelligence Scale	3
Background	3
Short Forms of the WISC	4
Testing Procedures	5
Findings	6
Verbal Development—Vocabulary Raw Scores	6
Performance—Block Design Raw Scores	8
Comparison With the WISC Standardization Group	8
Normalized Scale Scores	12
Construction of a Short Form	12
Percentile Equivalents	13
Discussion	13
Summary	16
References	18
List of Detailed Tables	20
Appendix I. Technical Notes	34
The Survey Design	34
Reliability	35
Extent of Missing Test Results and Imputation Procedures Used	36
Sampling and Measurement Error	37
Hypothesis Testing	38
Small Categories	39
Quality Control for the Psychological Test Battery	40

SYMBOLS

Data not available	. . .
Category not applicable	. . .
Quantity zero	-
Quantity more than 0 but less than 0.05	0.0
Figure does not meet standards of reliability or precision	*

INTELLECTUAL DEVELOPMENT OF YOUTHS

AS MEASURED BY A SHORT FORM OF THE WECHSLER INTELLIGENCE SCALE

James Scanlon, *Division of Health Examination Statistics*

INTRODUCTION

General

This report presents information on aspects of the intellectual development of youths 12 through 17 years of age in the noninstitutional population of the United States, as measured by two subtests of the Wechsler Intelligence Scale for Children (WISC) used in the Health Examination Survey of 1966-70. The two WISC subtests used, for reasons explained below, were the Vocabulary and Block Design subtests. Two reports describing WISC test results of a previous Health Examination Survey of children 6 through 11 years of age have been published.[1,2]

In this report on WISC findings from the survey of youths, analysis is limited to variations in test scores associated with age, sex, and grade placement. Construction of normalized scaled scores and of a short-form estimate of WISC Full Scale Scores for use in the analysis of other data collected in the survey is also described. Subsequent reports will explore the relationships between WISC scores and the behavioral, biomedical, socioeconomic, and other psychometric data gathered in the survey.

Source of the Data

Test results analyzed in this report were obtained from the Health Examination Survey, a major program of the National Center for Health Statistics (NCHS), authorized under the National Health Survey Act of 1956 and subsequent amendments as a continuing Public Health Service activity.[3] Complementing other NCHS programs aimed at assessing the health status of the American people, the Health Examination Survey collects and analyzes data gathered by direct physical examinations, tests, and measurements performed on probability samples of the U.S. population. The Survey is conducted as a series of separate, cross-sectional programs referred to as "cycles." Each cycle is limited to some specific segment of the U.S. population and to certain aspects of the health of that segment of the population. Since 1960, three separate surveys, or cycles, have been completed. Cycle I was concerned with adults aged 18 through 79 years in the noninstitutional population of the United States and was completed in 1962.[4,5] Cycle II was a survey of children aged 6 through 11 years, completed in 1965;[6,7] and Cycle III, which was completed in 1970,[8] focused on youths 12 through 17 years of age. For the survey of youths, on which this report is based, a probability sample of the Nation's noninstitutionalized youths 12 through 17 years of age was selected and examined. Field survey operations began in March 1966 and were completed in March 1970. Of the 7,514 youths selected for the sample, 6,768 were examined, a response rate of 90 percent. Because of the sample design, adjustment for nonresponse, and weighting procedures used, examination results

may be considered representative of the approximately 23 million noninstitutionalized youths in the United States 12 through 17 years of age at the time of the survey.[8] Sampling errors associated with estimates shown in this report are presented in appendix I.

Each youth in the sample was administered a 3-hour, single-visit examination in a mobile examination center specially constructed for the survey. The examination focused primarily on factors related to biological and psychological aspects of growth and development. It included examinations by a physician and a dentist, tests administered by a psychologist, and a variety of additional tests and measurements performed by technicians. To insure that the survey would provide comparable data on growth and development and on health characteristics throughout the continuum of childhood and adolescence, many of the tests and measurements performed on the youths are the same as those that were carried out on children aged 6 through 11 years in the previous cycle, with some modifications for the difference in ages. In addition, information relating specifically to adolescent health and behavior was collected.

To supplement data obtained from the examination, several questionnaires were employed. Among these were a household questionnaire administered by a U.S. Bureau of the Census interviewer to obtain demographic and socioeconomic information, two medical histories of the sample youth, one completed by a parent and another by the youth, and a health behavior questionnaire completed by the youth at the examination center. For those adolescents in school, information on grade placement, teachers' ratings of behavior and adjustment, and details of any health problems known to the teacher were requested from the school attended.

The grade placement of each sample youth was obtained from the questionnaire sent to the school each youth attended. For youths on summer vacation, the grade placement recorded was the grade the youth would enter in the fall. If the school questionnaire was not available for a sample youth, grade placement or the fact of having left school was obtained from the psychological test record forms.

Birth certificates were obtained for verification of the youth's age and for other facts relating to his or her birth. The age recorded for each youth was his age at last birthday as of the date of examination and was confirmed by comparison with the date of birth entered on the birth certificate. The age criterion for inclusion in the sample was defined in terms of age at the time of the first interview. Since the examination usually took place 2 to 4 weeks after the interview, some of the youths who were 17 years of age at the time of interview became 18 by the time of the examination. There were 58 such cases. In the adjustment and weighting procedures and in this analysis, these youths were included in the 17-year-old group.

All information was collected under conditions of confidentiality. More detailed information on the survey plan, sample design, examination content, and operation of the survey of youths is presented in appendix I and in a previous report.[8]

THE PSYCHOLOGICAL TEST BATTERY

The psychological test battery included in both the survey of children and the survey of adolescents was the result of consultation with child psychologists from the academic community and the National Institute of Mental Health. By using essentially the same battery of tests for youths as that used in the children's survey, it was believed possible to assess certain aspects of intellectual and, to some extent, emotional growth and development, on a comparable basis throughout childhood and adolescence.

The 70-minute battery contained measures of intellectual development (both verbal and performance), school achievement, literacy (both reading and writing), and aspects of personality development. The Vocabulary and Block Design subtests of the WISC and a modified version of the Goodenough-Harris Drawing Test were the principal measures of intellectual development used. Two subtests of the Wide Range Achievement Test (WRAT) were included to provide estimates of school achievement in the basic skills of oral reading and arithmetic computation. Inclusion of five cards of the Thematic Apperception Test (TAT) represented an attempt to measure some aspects of personality, as well as oral speech and communication. Finally, to assess the level of illiteracy among youths in the Nation, a brief test of literacy was developed

for the survey under a contract with the Educational Testing Service of Princeton, New Jersey, and administered to each youth in the sample. The development of this test has been described in an earlier report.[9]

A study evaluating the test battery as administered in the survey of children, which did not include the Brief Test of Literacy but was otherwise almost identical to the test battery employed in the present study, was conducted on a contractual basis by Dr. S. B. Sells of the Institute of Behavioral Research, Texas Christian University.[10] That study provided a review of the literature pertaining to research and evaluation of the battery components, recommendations concerning the types of inferences that could appropriately be made about the test results, and recommendations relating to additional research considered necessary to make proper use of the data collected.

An additional contractual study relating specifically to the WISC was completed by Jane Mercer and Joyce Smith of the University of California at Riverside.[11] This analysis evaluated the use of the Vocabulary and Block Design subtests as a basis for estimating Full Scale Scores among children from different socioeconomic levels and ethnic groups. It also examined the amount and direction of error likely to occur if these subtests were used to estimate rates of retarded intellectual development in those populations.

THE WECHSLER INTELLIGENCE SCALE

Background

The WISC, published in 1949, extended the well-known Wechsler scales for adolescents and adults into the age range of 5 through 15 years.[10,12] Since its publication, the WISC has been the subject of extensive investigation and has been used extensively in schools and clinics as an individual measure of general intellectual development.[10]

Much has been written about the concept of intelligence, and attempts to define it have been legion.[13] For the purposes of this report, it is necessary to focus on the concept of intelligence implicit in the Wechsler Scales. That concept, embodied in the WISC, assumes an aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment.[12,14,15] While intelligence is operationally defined and assessed in terms of performance on a composite of ability tasks, such as tests of verbal and spatial ability and tests of abstract and arithmetic reasoning, it is also considered to be part of the broader construct, personality.[12] The thinking underlying the WISC is that intelligence cannot be separated from the rest of the personality, and a deliberate attempt was made in the development of the test to take into account other personality factors, such as persistence, drive, and energy level, which were assumed to contribute to the total effective intelligence of the individual. This effort is reflected both in the composition of the WISC and in the equal weights assigned to the separate subtests.[12,14] No attempt was made in the WISC to construct a series of tests that measure "primary abilities" or to order them into a hierarchy of relative importance.

The WISC consists of 12 subtests (six Verbal and six Performance), two of which are treated as supplementary and not routinely used. An important innovation in the Wechsler scales was the use of the *deviation intelligence quotient.* This measure is considered a superior alternative to the *mental age* concept and evaluates test performance on the basis of distributions of scores of representative samples of persons of comparable chronological age. In the standardization of the WISC, Wechsler kept the standard deviation of intelligence quotients constant from year to year, with the result that when tested at different ages an individual's obtained deviation IQ would not vary unless his actual test performance varied when compared with that of his peers. *Raw scores* for each subtest are converted to *scaled scores.* These are normalized standard scores which have a mean of 10 and a standard deviation of 3 for each age level. The sum of the five scaled scores for the Verbal Series constitutes the Verbal Scale Score (VS); and similarly the Performance Scale Score (PS) is the sum of the scaled scores for the five subtests of the Performance Series. The Full Scale Score (FS) is the sum of the Verbal Scale and the Performance Scale Scores. Deviation intelligence quotients are provided in the WISC manual for VS, PS, and FS. Deviation IQ scales at each age

interval have a mean of 100 and a standard deviation of 15.

Short Forms of the WISC

Because time limitations must inevitably be imposed on a multidisciplinary endeavor such as the Health Examination Survey, no single factor, whether physiological, dental, physical, or psychological, could be evaluated as thoroughly as desired. Thus, compromises were necessary in planning the test battery. Time limitations did not permit administration of the full WISC in addition to the other tests and measurements selected for the appraisal of growth and development of adolescents. The test battery adopted provided for the administration of two subtests of the WISC, the Vocabulary and Block Design subtests. These were selected to serve as a short-form test from which estimates of Full Scale WISC IQ's could be made. In addition to providing a general index of intellectual development, the two subtests can be interpreted separately as measures of verbal development and the performance (nonverbal) aspects of intellectual development. Some discussion is appropriate, therefore, concerning the validity of a composite based on these subtests for the estimation of Full Scale Scores in the context of the Health Examination Survey.

A number of investigators have assessed the efficiency and accuracy of various combinations of two or more subtests of the WISC in estimating Full Scale Scores.[16-32] Many of these studies have been based on small samples of exceptional children such as the mentally retarded, the physically disabled, or children referred to child guidance clinics or social service agencies.

Correlations between a linear Vocabulary and Block Design combination and Full Scale Scores in these special populations have been reported by a number of investigators. Finley and Thompson[16] obtained a correlation of 0.68 among 309 mentally retarded children; Greenmun[17] obtained one of 0.92 among 632 referrals to the Child Guidance Clinic at Texas Christian University; Simpson and Bridges[18] found the correlation to be 0.87 for 120 children and youths referred to the Division of Child Guidance of the Oklahoma City Public Schools; Wight and Sandry[19] obtained 0.91 among 83 children hospitalized for a physical disability, and Mumpower[20] found the correlation to be 0.95 among 50 children and youths referred for evaluation to the Special Education Department at the University of Southwestern Louisiana.

One study, by Silverstein,[21] dealt with a fairly large sample of predominantly normal children. He determined the correlations with Full Scale Scores of all possible short forms of two, three, four, and five subtests using the WISC standardization data for 200 children at three age levels: 7½ years, 10½ years, and 13½ years. The best dyad predictor at age 10½ years was the Vocabulary and Block Design combination, while at ages 7½ and 13½ this dyad was among the better predictors. Differential weighting of the subtest scores did not result in appreciably higher predictive validities for Full Scale IQ.

A second large-scale study, mentioned above, was completed by Mercer and Smith under contract with the Health Examination Survey.[11] The authors used correlation and regression analysis of data from 1,310 children attending public elementary schools in Riverside, California, during the school year 1967-68 to evaluate the use of the Vocabulary and Block Design subtests of the WISC for predicting Full Scale Scores of children from different socioeconomic and ethnic groups. The sample consisted of children aged 6-11 years from three groups; 505 were English-speaking Caucasians (called "Anglos" in the study), 487 were of Mexican-American heritage; and 318 were Negro. Of the total sample, 1,270 were enrolled in regular classes and 40 were enrolled in classes for the educable mentally retarded. When Vocabulary and Block Design scaled scores were used to predict Full Scale, Verbal, and Performance IQ's, it was found that estimates were not materially improved by categorizing the separate ethnic groups according to socioeconomic status or age. Multiple correlation coefficients were approximately the same for all three ethnic groups, and the authors concluded that Vocabulary and Block Design provided an essentially equivalent short form for predicting Full Scale Scores for the three ethnic groups. The multiple correlation coefficients were 0.87 for Anglo

children, 0.83 for Negro children, and 0.85 for those of Mexican-American heritage; the standard errors of estimate were 7.5, 6.8, and 6.7, respectively.

A score representing the simple sum of the two subtest scaled scores yielded a correlation of 0.88 with Full Scale Scores for all ethnic groups combined. Vocabulary and Block Design were the optimal dyad for Anglo and Mexican-American children. Vocabulary and Object Assembly were the optimal dyad for Negro children, but the difference in accuracy of prediction was minimal (R = 0.835 versus 0.827, or a 0.008 difference). As expected, Vocabulary Scale Scores in this study correlated quite highly with Verbal Scale Scores for all three ethnic groups (0.82 for Anglo children, 0.80 for Negro children, and 0.82 for Mexican-American children). Similarly, Block Design Scale Scores were found to correlate 0.76 with Performance IQ among Anglo children, 0.70 among Negro children, and 0.71 among Mexican-American children. The authors concluded that no other dyad would have produced better overall predictions of Full Scale IQ for the three ethnic groups in their sample, and that the Full Scale IQ for that sample could be predicted with about the same accuracy from the unweighted sum of the two scaled scores as from differentially weighted scaled scores.

Hite has confirmed Wechsler's data indicating that Vocabulary and Block Design are the most reliable subtests of the WISC battery.[12,33] Hagen and Cohen in the United States and Gault in Australia have reported that both of these subtests are highly loaded on the so-called "general" factor of intelligence obtained in factor analytic studies of the WISC over the entire age range of 5 through 15 years.[34-36]

Hence the weight of all available evidence supports the conclusion that the two subtests selected for use in the survey constitute the best dyad available for estimating Full Scale IQ and that the validity of this dyad for the estimation of Full Scale Scores is high. While this dyad is valid and appropriate for the analysis of group statistics in a research setting such as the Health Examination Survey, its use as the sole instrument in the comprehensive assessment of intellectual functioning in clinical settings is obviously not advocated.[11]

TESTING PROCEDURES

The Vocabulary and Block Design subtests of the WISC were the second and third procedures in the 70-minute psychological test battery, following administration of the Wide Range Achievement Test. The testing was accomplished in a small, adequately lighted, climate-controlled, sound-treated room in the mobile examination center. The psychologists who administered these tests had at least a master's degree and had experience in administering tests. There were two psychologists on the examining team at all times. The examiners were selected, trained in the field-testing procedures, and supervised by the Office of the Psychological Advisor to the Health Examination Survey, under the direction of an experienced Ph.D.-level psychologist. In the initial training and ensuing supervision of the examiners, strong emphasis was placed on uniform methods of test administration, scoring, and recording of data. During the course of the survey of youths, a total of 12 examiners participated in administration of the battery. Quality-control procedures employed in administration of the psychological test battery are described in appendix I.

The standard WISC Record Form was used for recording responses and scores. The subtests were administered as specified in the WISC Manual except that, in most cases, each youth was started with the sixth word on the Vocabulary subtest.

The Vocabulary subtest of the WISC consists of 40 words arranged in order of increasing difficulty. The youth was asked what each word means. The response to each word was scored "2," "1," or "0," except words 1 through 5 which were scored "2" or "0." The maximum raw score was 80. The test was discontinued after five consecutive failures. Since each youth normally started at word 6, he was automatically credited with a score of 2 for words 1 through 5 if 2-point responses were given for words 6 through 10. If no credit or a 1-point response was given for any of words 6 through 10, the examiner went back to word 5 and worked backward until the criterion of five consecutive 2-point responses was met, or all of words 1 through 5 had been administered.

The Block Design subtest of the WISC requires the examinee to reproduce designs constructed of small colored wooden cubes within prescribed time limits. There are 10 such designs, increasing in complexity and requiring either four or nine cubes. Administration and scoring followed procedures outlined in the test manual. The maximum score was 55 points.

A further word on the sample design and its possible effect on WISC findings in this report is appropriate. Because of both operational considerations and the desire to obtain longitudinal data, over 30 percent of the youths comprising the Cycle III sample had also participated in the survey of children completed in 1965. Since the shortest test-retest period in the two surveys was about 2½ years, it was assumed that WISC results for the retested youths were not appreciably influenced by any practice effect.

FINDINGS

Verbal Development—Vocabulary Raw Scores

Variations with age.—As mentioned above, the Vocabulary subtest of the WISC is a measure of verbal aspects of intellectual development, correlating highly with both the Verbal Scale and Full Scale Scores of the WISC. Results of this test from the Health Examination Survey indicate, not surprisingly, that verbal development is positively correlated with age within the age group of 12 through 17 years (table 1 and figure 1). Mean Vocabulary raw scores increased from 36.5 points among 12-year-olds to 45.2 points among 17 year-olds. The highest possible raw score was 80 points. Examination of unsmoothed and smoothed mean Vocabulary raw scores for 4-month age groups reveals a similar

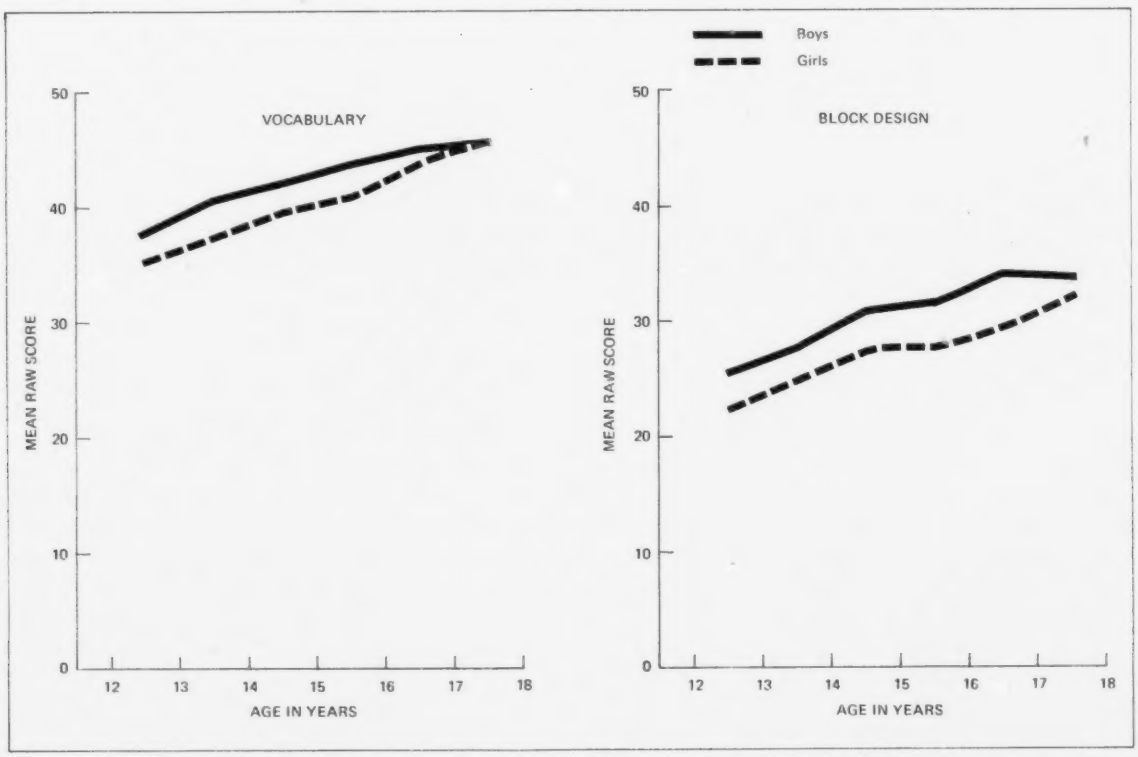

Figure 1. Mean raw scores on the Vocabulary and Block Design subtests of the WISC, by sex and age: United States, 1966-70.

trend of higher mean raw scores with age, but at a decelerating rate (table 2 and figure 2). The curves were smoothed using a 3-point moving average technique.

The variability in raw scores attained by the youths, as indicated by the standard deviations of the Vocabulary subtest raw score distributions at each 1-year age level, tended to increase very slightly with age (table 1 and figure 3). An approximate index of relative variation, the coefficient of variation (ratio of the standard deviation to the mean) remained fairly constant for the Vocabulary subtest scores over all age groups. The two highest raw scores found in this national sample, scores of 76 and 73, were attained by two 16-year-olds. Thus, it appears that the subtest had an adequate "ceiling" for the age group tested in the survey.

Sex differences in scores.—Boys, in general, tended to perform slightly better than girls on the Vocabulary subtest (table 1 and figure 1). At every 1-year age level except the oldest, the mean score for boys exceeded that for girls. The differences ranged from 1 to 3 raw score points, but were large enough to be statistically significant, at the 5-percent confidence level only for the 12-, 13-, and 15-year-olds. At age 17, the means for boys and girls were identical. (Hypothesis testing procedures are described in appendix I.)

Grade placement.—As expected, mean vocabulary raw scores tended to increase with successive grade levels (table 1 and figure 4). This grade differentiation was true for all age groups combined and for each 1-year age group. The mean vocabulary raw score among those

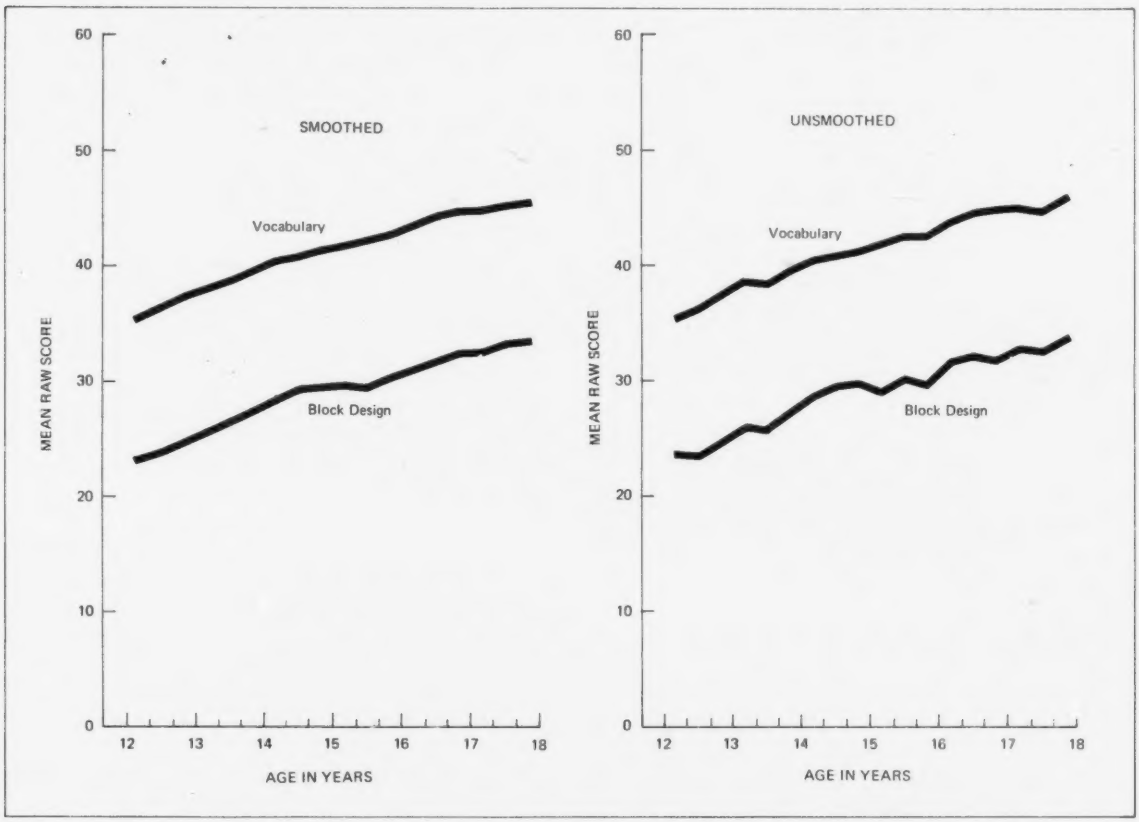

Figure 2. Unsmoothed and smoothed mean Vocabulary and Block Design raw scores on the WISC, by age in 4-month age intervals: United States, 1966-70.

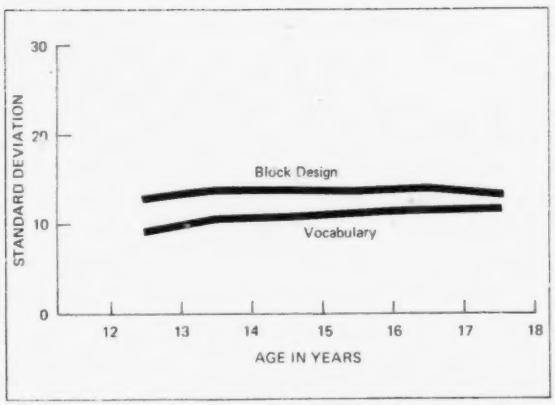

Figure 3. Standard deviations of raw scores on the Vocabulary and Block Design subtests of the WISC, by age: United States, 1966-70.

relatively few youths, mostly 12- and 13-year-olds, whose grade placement at the time of testing was less than 5th grade was 18.9, while the mean for youths in the 12th grade at the time of testing was 49.0. High school graduates with no additional formal schooling had a mean of 47.5, and youths who were attending school beyond high school had a mean of 54.1. Youths who had left school before graduating (mostly 16- and 17-year-olds) attained a mean of 32.1. Generally, in each grade younger individuals obtained higher mean raw scores than did older ones (table 1).

Consistent with the sex differences in mean scores within each age group reported above, boys attained higher mean scores than did girls at almost every grade level, including youths in special ungraded classes for the mentally retarded and physically handicapped, and those who left school before graduating.

Performance—Block Design Raw Scores

Variations with age.—The Block Design subtest is an approximate measure of the performance or nonverbal aspects of intellectual development. Results for this subtest also indicate a positive association with age in the 12- through 17-year-old population studied, with a tendency to level off in the oldest groups (figures 1 and 2). Mean raw scores on the Block Design subtest increased from 24.0 among 12-year-olds to 33.0 among 17-year-olds (table 3). The highest possible raw score on this test was 55 and was attained by about 0.4 percent of the sample youths.

Variability in Block Design raw scores was higher at every age level and for both sexes than that observed for the Vocabulary subtest (tables 1, 2, and 3 and figure 3). The larger Block Design standard deviations, however, are probably a result of the 0 versus 4 scoring of most items on this subtest, not counting bonus points. No age trends in standard deviations were apparent, and the relative variation of Block Design raw scores tended to decrease somewhat with increasing age.

Sex differences in scores.—Throughout the age range tested, boys on the average scored higher than did girls on the Block Design subtest (table 3). The differences were statistically significant at all age levels and ranged from about two to four raw score points.

Grade placement.—On the Block Design subtest, as on the Vocabulary subtest, mean raw scores tended to increase with successive grade levels among both boys and girls (table 3). This was true for each 1-year age group. At every grade level, boys attained higher scores than girls, including youths in special placement and those who left school before graduating. The sex differences in mean scores were large enough to be statistically significant at the 5-percent confidence level for youths in the 6th through the 12th grades and for those in special placement. There was a tendency for younger children in each grade to perform somewhat better than did older children on this test, up to about the 10th grade (table 3).

Comparison With the WISC Standardization Group

The WISC was standardized in 1949 on a sample of 100 boys and 100 girls at each year of age from 5 through 15.[12,37] The standardization sample also included 55 children, most of them institutionalized, who were in the chronological age range 5 through 15 years and determined by staff psychologists of their respective institutions or schools as having IQ's under 70 but not below 50. With the exception

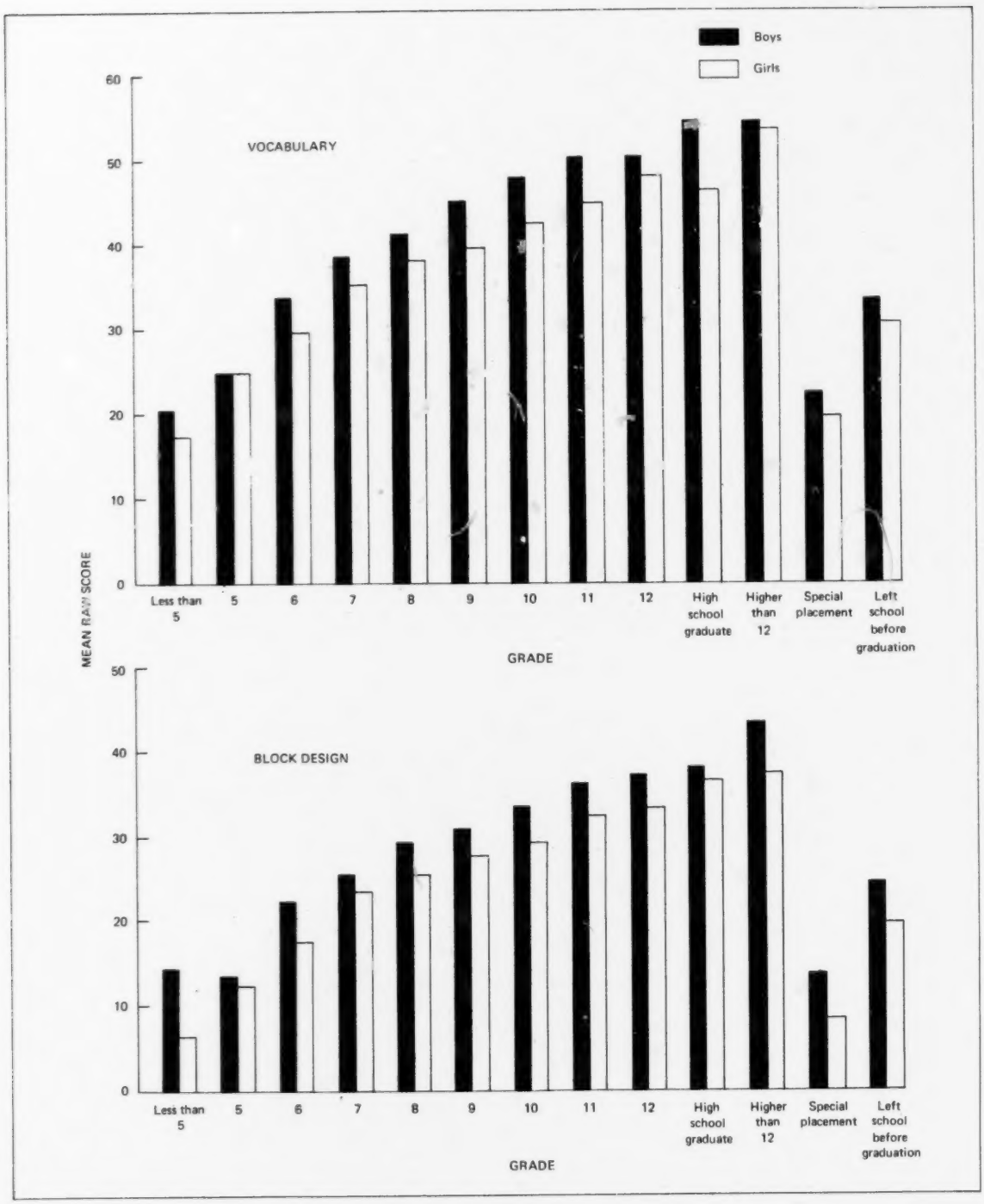

Figure 4. Mean raw scores on the Vocabulary and Block Design subtests of the WISC, by grade in school and sex: United States, 1966-70.

of the known low IQ group, most of whom were tested within 2 months of their midyear, each child was tested within 1½ months of his midyear. Only white children were included in the standardization group.

Within each year of age and for the total group, the standardization sample was selected so that it would be representative of the total U.S. population, according to data from the 1940 U.S. Census of Population, with regard to the following three variables:

1. The proportion of the U.S. population living in each of four geographic regions of the country—(1) the region comprised of the New England and Middle Atlantic States, (2) the region comprised of the North Central States, (3) the region comprised of the South Atlantic and South Central States, and (4) the region comprised of the Mountain and Pacific States.
2. The proportion of the 1940 U.S. population living in rural and urban areas.
3. The proportion of the total adult employed white male population in nine occupational classification groups.

The Midwest group was reduced in size and the Western group was made slightly larger than the U.S. proportions in 1940 to allow for wartime and postwar population shifts during the 1940's. Standardization testing took place in 85 communities. The sample was drawn from school populations, except for the majority of the 55 children of known low IQ's, most of whom were institutionalized.

Except for a slight bias for urban and small-town areas as opposed to rural areas, and aside from the fact that the standardization was accomplished on a native white population only, the standardization sample for the WISC has been regarded as good.[10]

The present study was based on findings from 6,768 examined youths who may be considered representative of the estimated 23 million noninstitutionalized youths 12-17 years old in the United States as of the midpoint of the survey. Testing took place at 40 locations between 1966 and 1970. The sample design used in the survey was a multistage, stratified probability sample of households in land-based segments. A more detailed description of the sampling design is presented in appendix I and in previous publications.[6-8] The sampling frame insured the representativeness of the sample with respect to degree of urbanization, region, and the rate of population change from 1950 to 1960, the latter being an indirect indicator of the economic condition of the area. Data used in this report for each sample youth were inflated in the estimation process to characterize the larger universe of which the sample youths are representative and to include an adjustment for the nonresponse group. This made the final sample estimates of the populations agree exactly with independent controls from the Bureau of the Census for the U.S. noninstitutionalized population as of the approximate midpoint of the survey, with regard to age in years, race, and sex for the study population. Thus, findings from the present study are based on from approximately 960 to 1,200 youths at each year of age 12 through 17, and may be considered representative of the total noninstitutionalized U.S. population of 12- through 17-year-olds with regard to age, race, sex, region, population size, and rate of population change between 1950 and 1960.

Comparisons of raw score means and standard deviations on the Vocabulary and Block Design subtests for the 1949 WISC standardization sample and U.S. estimates from the national Health Examination Survey are shown in figures 5 and 6. WISC standardization sample means and standard deviations were estimated from the conversion tables in the WISC Manual. For every 4-month age group in the range 12 through 15 years, estimated WISC standardization means for the Vocabulary subtest exceeded unsmoothed means from the Health Examination Survey. The differences ranged from about 3 to 6 raw score points. For the Block Design subtest also, estimated WISC standardization group means ranged from about 2 to 8 raw score points higher than corresponding Health Examination Survey findings.

Using data from the previous survey of 6- to 11-year-olds, it is possible to compare Vocabulary and Block Design subtest characteristics from the original Wechsler standardization

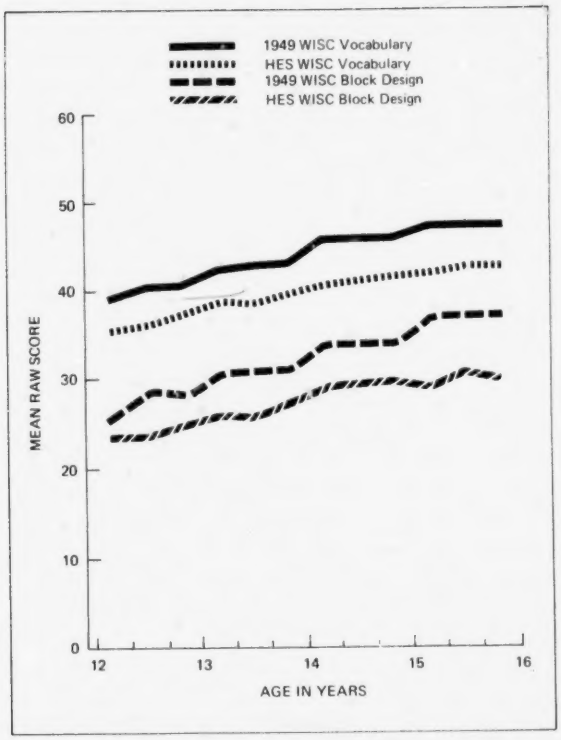

Figure 5. Mean raw scores on the Vocabulary and Block Design subtests of the WISC for the WISC standardization group and for U.S. estimates from the Health Examination Survey, at 4-month age intervals.

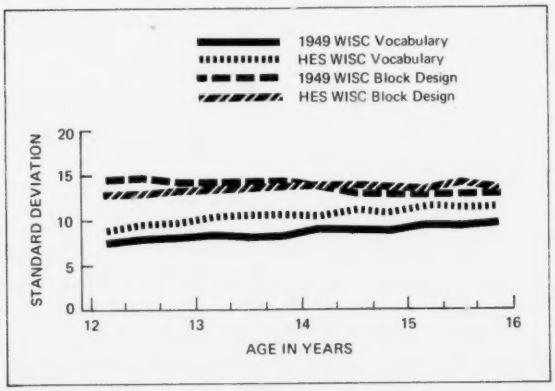

Figure 6. Standard deviations of raw score distributions on the Vocabulary and Block Design subtests of the WISC for the WISC standardization group and for U.S. estimates from the Health Examination Survey, at 4-month age intervals.

sample and the national Health Examination Surveys in more detail. The comparisons are made for ages 7½, 10½, and 13½, ages which the test developers felt were probably most representative of the age range for which the WISC was designed.[1,2] At the 7½-year age level, the mean Vocabulary subtest raw score was lower in the Cycle II sample of the Health Examination Survey than that estimated from the original WISC sample, and the variability was greater. Both differences were statistically significant.[1] At age 10½ years, the mean and variability of the distribution of Vocabulary raw scores were both slightly, but not significantly, greater in the Health Examination Survey sample than in the WISC sample.[1] Among the 13½-year-olds, the Health Examination Survey mean Vocabulary raw score was significantly lower than the original standardization figure, while the variability was slightly, but not significantly greater.

For the Block Design subtest, mean raw scores in the Health Examination Survey exceeded those found in the WISC standardization sample at ages 7½ and 10½ and were lower than those found in the WISC group at age 13½. The differences were statistically significant for the 7½- and 13½-year age groups. Variability of the distribution of Block Design raw scores was slightly greater in the Health Examination Survey sample than the WISC sample at age 7½, significantly smaller than the WISC group at age 10½, and slightly, but not significantly, smaller than the WISC sample at age 13½.

The correlations between scaled scores on the Vocabulary and Block Design subtests in the national Health Examination Survey are in fairly close agreement with those of the original standardization sample. At age 7½, the Pearson Product-Moment coefficient of correlation found in the Health Examination Survey, 0.38 ± 0.027, was slightly higher than the standardization sample figure of 0.33, while at age 10½ it was slightly lower, 0.50 ± 0.022, compared with 0.54 in the WISC group. At age 13½, the correlation coefficient found in the Health Examination Survey was 0.54 ± 0.159, somewhat higher than that in the standardization sample, which was 0.42.

Boys tended to score higher than girls on the verbal and performance measures in both the original WISC standardization group, which

employed all 12 subtests of the WISC battery, and in the present study, which was limited to one verbal subtest and one performance subtest. Seashore, Wesman, and Doppelt found that, on the average, boys in the standardization group did slightly better than girls at each year of age from 12 through 15.[37] The differences found in that study were small, but consistent, occurring primarily in the older age groups. On the total Verbal Scale, boys excelled girls by more than 3 points at age 8 and ages 10 through 15. On the full Performance Scale, the differences favored boys by more than 3 points at ages 8 and 10, while the girls did better at ages 5, 6, 7, and 9.[36]

In the Health Examination Survey, which included 16- and 17-year-olds, boys also tended to do slightly better on the tests than did girls. On the Vocabulary subtest, average scores for boys exceeded those for girls by 1 to 3 raw score points at every 1-year age level except the oldest, when the means were identical. The differences were statistically significant, however, only among the 12-, 13-, and 15-year-olds. On the Block Design subtest, boys, on the average, outscored girls at every age level from 12 through 17. The differences were statistically significant at all age levels and ranged from about 2 to about 4 raw score points.

Normalized Scaled Scores

To express scores in a form that indicates a youth's performance on a test relative to his age group, to make such scores comparable among different age groups, and to construct a common metric for the comparison of two or more tests or subtests of different length, raw scores must be converted to some common measure. In the original WISC standardization process, normalized standard scores with a mean of 10 and a standard deviation of 3 were constructed for each of the 12 subtests within each age level. The WISC Manual provides such scores for each 4-month age interval from age 5 through 15 years.

In the present study raw scores from the total U.S. sample were similarly converted to normalized standard scores with a mean of 10 and a standard deviation of 3 for each subtest within each 4-month age interval.[38] As in the WISC standardization process, minor irregularities in the progression of scaled score equivalents from age to age were smoothed. It was assumed that these irregularities were due to sampling variability. Because there were few instances of such minor deviations, and these were at the extremes of the raw score distribution, the scaled scores are essentially a direct translation from raw scores to a normalized distribution of scaled scores with a mean of 10 points and a standard deviation of 3 points. Tables 6 and 7 present the scaled score equivalent of each raw score from the present study.

Mean scaled scores on the Vocabulary and Block Design subtests are shown according to age in years, sex, and grade placement in tables 4 and 5 and parallel the findings on raw scores reported above. Mean scores on each subtest were generally slightly higher for boys than for girls and tended to increase with successive grade levels. Younger children in each grade tended to attain higher scaled scores than did older children in the same grade. Among the relatively small number of youths in the sample who left school before graduating, both boys and girls obtained higher mean scores on the Block Design subtest than on the Vocabulary subtest; the differences were statistically significant for both boys and girls.

Construction of a Short Form

To carry out analyses of the interrelationships among intellectual development and other survey variables, some approximate index of general intellectual development is required, incorporating both the Vocabulary and Block Design subtests. For this purpose it was decided to employ a linear transformation of the simple sum of the scaled scores on the two subtests into a new distribution of standard scores with a mean of 100 points and a standard deviation of 15. This method weights the two subtests equally and makes use of the recent representative national norms provided for each subtest by the survey. The resulting standard scores are of the same order of magnitude as the WISC Full Scale IQ; and as constructed in the present study, can range from 43 to 157. Such an index is obviously not as reliable as one based on all 10 subtests. It is, however, a reasonable estimate of intellectual development, suitable

for the research purposes of the survey.[10,11] The basis for converting the sum of the scaled scores into index scores in this study is shown in table 8.

Variations in mean index scores associated with age, sex, and grade placement are presented in table 9. With the two subtest results combined, boys again tended to outscore girls, the mean differences being statistically significant for each 1-year age group except for the 17-year-olds.

Variations in mean scores with grade are similar to those observed among the separate subtest scaled scores. Mean scores tended to increase with successive grade levels. Youths in the normal, or modal, grade placement for their age obtained mean scores slightly above 100, the average score. Youths who were accelerated in grade placement tended to obtain above-average scores, while older youths in each grade obtained lower than average scores.

Percentile Equivalents

Another index of a youth's relative standing in his age group with regard to a particular characteristic is the percentile equivalent. Percentile equivalents, as defined in this report, represent the percentage of individuals in the sample falling below particular raw scores. Table 10 presents selected percentile equivalents of raw scores for both subtests, based on results obtained in the Health Examination Survey. The percentile equivalents were computed at 1-year age intervals from 12 through 17 for both sexes combined and for boys and girls separately and were smoothed at the extremes of the distributions. Thus, a 14-year-old obtaining a raw score of 48 on the Vocabulary subtest in this national sample surpassed 75 percent of the 14-year-olds tested. Unsmoothed cumulative percent distributions of raw scores for each subtest, plotted on normal probability paper, are shown in figure 7. Cumulative percentages derived from a normal distribution would yield a straight line curve on such a graph. Thus it appears that Vocabulary subtest raw scores are approximately normally distributed within each age group in this study. The cumulative percent distribution of Block Design raw scores is much more irregular, particularly at the lower raw score levels.

DISCUSSION

Though the roots of testing itself are lost in antiquity, the major development shaping the contemporary testing of general intellectual level had its effective beginning with the development of the Binet-Simon Scale in 1905. Binet and his coworkers were seeking to measure general intelligence in an educational setting in order to identify children in Paris schools who would require special educational services.[39,40] The first such tests were considered successful because they differentiated between children known to do well in school and those known to perform poorly in the classroom. Most later tests of intellectual level were patterned after and validated following the Binet model, thus continuing the influence of the scholastic aptitude orientation on intelligence testing.[39] These tests are thus indicators of both previous opportunity to learn and of ability to learn something new in a scholastic setting.

Despite the fact that Wechsler developed the WISC in protest against the measurement concept of mental age implicit in the Stanford revision of the Binet Scales and despite several important differences between the WISC and the Stanford-Binet, a number of WISC validation studies have used the Stanford-Binet as a criterion. Correlations between WISC Full Scale Scores and Stanford-Binet IQ's tend to be high, generally in the mid-80's, with some as high as the reliability coefficients of the tests.[10,39]

A number of investigators have found concurrent validity coefficients between WISC scores and achievement tests or other academic criteria clustering around 0.60.[39] As pointed out above, the Vocabulary and Block Design subtests of the WISC are the most reliable of the entire Scale, and a short form consisting of this dyad appears to be about the best available, given the constraints and research purposes of the Health Examination Survey.

This study makes available for the first time Vocabulary and Block Design subtest findings for a national probability sample of the noninstitutionalized population of youths 12 through 17 years of age with proportionate representation from all races. The sample is several times larger than that used in the original test standardization and hence should provide an even more

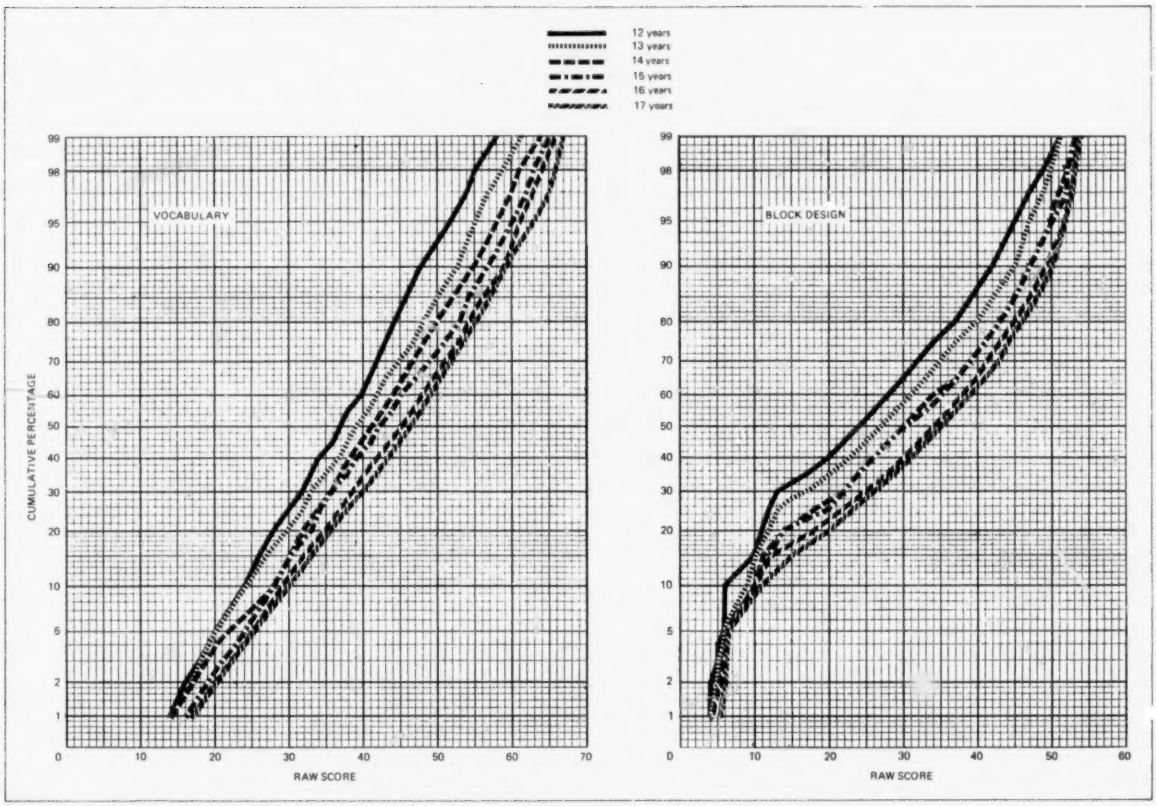

Figure 7. Cumulative percent distributions of WISC Vocabulary and Block Design raw scores by age, plotted on normal probability paper: United States, 1966-70.

stable base for standardization of the two subtests than did the original standardization sample.

The finding from the present study that Health Examination Survey raw score means were somewhat lower than those estimated from the WISC standardization group in the age range 12-15 years is interesting. Though this study can provide no definitive explanation, one likely reason is the difference in the two sample designs. As mentioned earlier, Wechsler's standardization was carried out in the late 1940's on a sample of white persons only, although it did include some children who were institutionalized. Health Examination Survey findings are based on a national probability sample designed to be representative of the total, noninstitutional population aged 12 through 17 years in the United States during the period 1966-70, with respect to age, sex, race, region, population size, and other socioeconomic variables.

A second finding from the present study that boys tend to outscore girls on at least these two subtests of the WISC is consistent with other analyses, including the Health Examination Survey of children 6 through 11 years of age,[1]

Mercer and Smith,[11] and Seashore et al., in their original standardization study of the WISC.[37] Regarding these differences in scores between boys and girls, Seashore and his coworkers concluded that the most reasonable assumption was that boys and girls are equal in levels of mental ability, but either the test items chosen turned out to be slightly biased in favor of boys or the 1949 standardization sample of boys was somehow biased; or that both biases may be responsible for the differences. This conclusion was based on the fact that Terman and Merrill found similar results in their 1937 revision of the Stanford-Binet Scales and could also find no definitive explanation from their data.[41] Findings from the present study and the previous survey of children, however, are based on large samples that are closely representative of the total U.S. noninstitutionalized population of 6-17-year-olds. In the light of the results of these two HES investigations, the hypothesis of bias in the standardization sample selection of children and youths does not explain the sex differences in scores. Although Seashore's data did not show test results separately for the two subtests used here, it is likely that the persistent difference between the scores of boys and girls on these subtests would have been found in the original study because of the high correlation between these two subtests and the total Verbal and Performance Scales of the WISC.

Another implication of the WISC results obtained in Cycles II and III can be drawn from the variation in mean indexes of verbal and performance level over the age range 6-17 years. Using raw score data from both cycles it is possible to construct approximate group growth curves for verbal development, as measured by the Vocabulary subtest, and for nonverbal aspects of intellectual development, as assessed by the Block Design subtest. These curves are not true growth curves because the approach here was essentially cross sectional rather than longitudinal. Also, an interval of about 3 calendar years separated the midpoints of the two surveys, and as mentioned earlier, some of the children examined in Cycle II had become eligible for and were selected for inclusion as sample youths in Cycle III, thus introducing a quasi-longitudinal aspect to these two sequential cross-sectional surveys. Because the time interval between the two surveys was small, it is likely that test scores from the two surveys are comparable and can be treated, for the most part, as resulting from a cross-sectional design. In addition, the fact that the age range encompassed in the two surveys was relatively small tends to minimize one of the major difficulties in the analysis of growth data from cross-sectional designs, namely, the differential effect that major cultural changes may have on the population characteristics of some but not all age groups studied.[13,42]

Smoothed Vocabulary and Block Design raw score means for each 4-month age interval from ages 6 through 17 are shown in figure 8. The curves were smoothed using a 3-point moving average technique. The curves of both Vocabulary and Block Design raw score means illustrate a rapid rate of increase in the younger groups and a somewhat decelerating rate of increase in the older age ranges. Unsmoothed means for each 1-year age group from 6 through 17 years of age appear in figure 9 and generally illustrate the same trend. Notably absent from these curves and from curves of changes in smoothed raw score means from age to age is the "adolescent growth spurt," a phenomenon typical of many physical aspects of growth and development.[43] Recent longitudinal studies of intellectual growth suggest that intelligence continues to increase gradually after the teens, that many of the more complex intellectual skills reach their peak as late as age 50, and that only items that depend on speed and intensity of response show the decline in the mid-twenties long considered characteristic of intellectual development.[13]

A more detailed discussion of sex differences in scores on the Wechsler scales and of changes in scores over age has been published.[15]

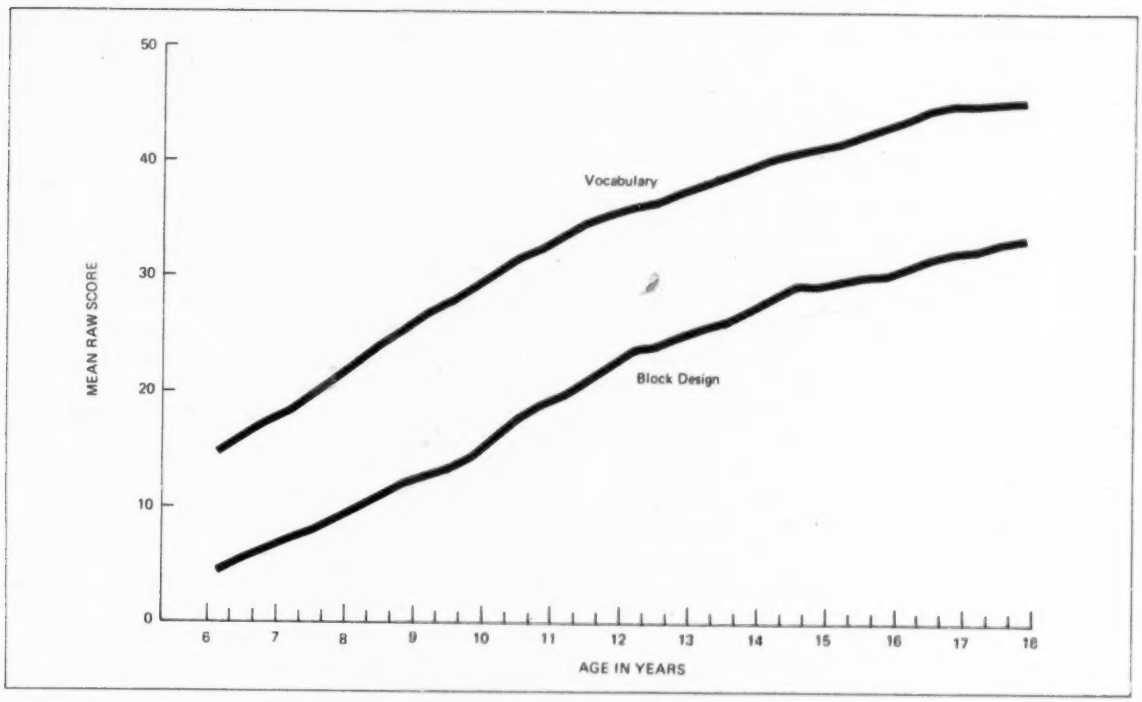

Figure 8. Smoothed mean raw scores on the Vocabulary and Block Design subtests of the WISC, by age in 4-month intervals: Cycles II and III of the Health Examination Survey.

SUMMARY

This report has presented findings on two subtests of the WISC, Vocabulary and Block Design, administered to a national probability sample of youths in the Health Examination Survey of 1966-70. In the survey, a probability sample of 7,514 youths was selected to represent the 23 million noninstitutionalized youths, 12 through 17 years of age, in the United States. Of the 7,514 youths selected in the sample, 6,768, or 90 percent, participated in the survey. Because of the sample design, adjustment for nonresponse, and weighting procedures used in the survey, findings for these youths may be considered to be representative of the total noninstitutionalized population of 12- through 17-year-olds in the United States with respect to age, sex, race, region, and other socioeconomic characteristics.

Variations in raw scores and scaled scores associated with age, sex, and grade placement have been discussed in this report. Percentile equivalents have been constructed, and the derivation of normalized scaled scores and a short-form estimate of WISC Full Scale Scores for subsequent use in analysis of survey data have been described.

Findings on the two subtests have been compared with data estimated from Wechsler's

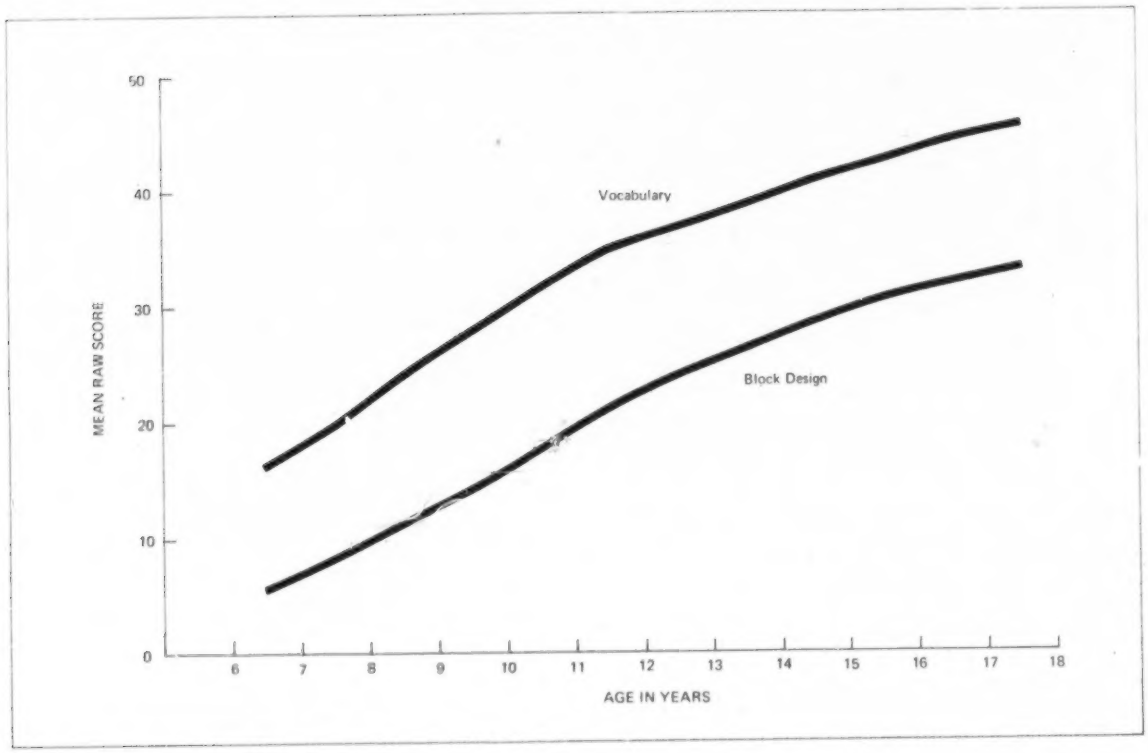

Figure 9. Mean raw scores on the Vocabulary and Block Design subtests of the WISC, by age: Cycles II and III of the Health Examination Survey.

original standardization sample, and variations in mean subtest raw scores over the age range 6-17 years have also been discussed.

Selected findings include the following relationships:

1. Mean raw scores on both subtests tended to increase with age, with a tendency to level off in the older age groups.

2. For every 4-month age group in the 12-15-year age range, Health Examination Survey mean raw scores on each subtest were lower than those found in the original standardization sample, which consisted of white youths only but included some children and youths in institutions.

3. Mean raw scores for boys generally exceeded those for girls.

4. Scaled Scores and estimated Full Scale Scores exhibited a tendency to increase with successive grade levels.

5. Youths who had left school before graduating tended to score significantly higher on the nonverbal test than on the verbal test.

REFERENCES

[1]National Center for Health Statistics: Intellectual development of children as measured by the Wechsler Intelligence Scale, United States. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 11-No. 107. Public Health Service. Washington. U.S. Government Printing Office, Aug. 1971.

[2]National Center for Health Statistics: Intellectual development of children by demographic and socioeconomic factors, United States. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 11-No. 110. Public Health Service. Washington. U.S. Government Printing Office, Dec. 1971.

[3]National Center for Health Statistics: Origin, program, and operation of the U.S. National Health Survey. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 1-No. 1. Public Health Service. Washington. U.S. Government Printing Office, Apr. 1965.

[4]National Center for Health Statistics: Plan and initial program of the Health Examination Survey. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 1-No. 4. Public Health Service. Washington. U.S. Government Printing Office, July 1965.

[5]National Center for Health Statistics: Cycle I of the Health Examination Survey: Sample and response, United States, 1960-1962. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 11-No. 1. Public Health Service. Washington. U.S. Government Printing Office, Apr. 1964.

[6]National Center for Health Statistics: Plan, operation, and response results of a program of children's examinations. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 1-No. 5. Public Health Service. Washington. U.S. Government Printing Office, Oct. 1967.

[7]National Center for Health Statistics: Sample design and estimation procedures for a National Health Examination Survey of children. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 43. Public Health Service. Washington. U.S. Government Printing Office, Aug. 1971.

[8]National Center for Health Statistics: Plan and operation of a Health Examination Survey of U.S. youths, 12-17 years of age. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 1-No. 8. Public Health Service. Washington. U.S. Government Printing Office, Sept. 1969.

[9]National Center for Health Statistics: Development of the Brief Test of Literacy. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 27. Public Health Service. Washington. U.S. Government Printing Office, Mar. 1968.

[10]National Center for Health Statistics: Evaluation of psychological measures used in the Health Examination Survey of children ages 6-11. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 15. Public Health Service. Washington. U.S. Government Printing Office, Mar. 1966.

[11]National Center for Health Statistics: Subtest estimates of the WISC Full Scale IQ's for children. *Vital and Health Statistics.* Series 2, No. 47. DHEW Pub. No. (HSM) 72-1047. Washington. U.S. Government Printing Office, Mar. 1972.

[12]Wechsler, D.: *Wechsler Intelligence Scale for Children, Manual.* New York. The Psychological Corporation, 1949.

[13]McCandless, B. R.: *Adolescents: Behavior and Development.* Hinsdale, Illinois. The Dryden Press Inc., 1970.

[14]Wechsler, D.: *The Measurement and Appraisal of Adult Intelligence* (4th ed.). Baltimore. Williams and Wilkins, 1958.

[15]Matarazzo, J. D.: *Wechsler's Measurement and Appraisal of Adult Intelligence* (5th ed.). Baltimore. The Williams and Wilkins Company, 1972.

[16]Finley, C. J., and Thompson, J.: An abbreviated Wechsler Intelligence Scale for children for use with the educable mentally retarded. *Am.J.Ment.Deficiency* 63:473-480, 1958.

[17]Greenmun, R.: *Abbreviated Forms of the Wechsler Intelligence Scale for Children.* Unpublished master's thesis. Texas Christian University, Aug. 1965.

[18]Simpson, W. H., and Bridges, C. C., Jr.: A short form of the Wechsler Intelligence Scale for Children. *J.Clin.Psychol.* 15(4):424, Oct. 1959.

[19]Wight, B., and Sandry, M.: A short form of the Wechsler Intelligence Scale for Children. *J.Clin.Psychol.* 18(2):166, Apr. 1962.

[20]Mumpower, D. L.: The fallacy of the short form. *J.Clin.Psychol.* 20(1):111-113, Jan. 1964.

[21]Silverstein, A. B.: Validity of WISC short forms at three age levels. *J.Consult.Psychol.* 31:635-636, 1967.

[22]Enburg, R., Rowley, V. N., and Stone, B.: Short forms of the WISC for use with emotionally disturbed children. *J.Clin. Psychol.* 17(3):280-284, July 1961.

[23]McKerracher, D. W., and Watson, R. A.: Validation of a short form WISC with clinic children. *Brit.J.Educ.Psychol.* 38:205-208, 1968.

[24]Nickols, J., and Nickols, M.: Brief forms of the WISC for research. *J.Clin.Psychol.* 19(4):425, Oct. 1963.

[25]Schwartz, L., and Levitt, E. E.: Short forms of the Wechsler Intelligence Scale for Children in the educable, non-institutionalized mentally retarded. *J.Educ.Psychol.* 51(4):187-190, 1960.

[26]Thompson, J. M., and Finley, C. J.: An abbreviated WISC for use with gifted elementary school children. *California J.Educ.Res.* 14(4):167-177, Sept. 1963.

[27]Yalowitz, J. M., and Armstrong, R. G.: Validity of short forms of the Wechsler Intelligence Scale for Children. *J.Clin. Psychol.* 11(3):275-277, July 1955.

[28]Silverstein, A. B.: Estimating Full Scale IQ's from WISC short forms. *Psychol.Rep.* 20:1264, June 1967, part 2.

[29]Silverstein, A. B.: A short form of Wechsler's scales for screening purposes. *Psychol.Rep.* 21:842, 1967.

[30]Silverstein, A. B.: A short form of the WISC and WAIS for screening purposes. *Psychol.Rep.* 20:682, Apr. 1967.

[31]Silverstein, A. B.: Reappraisal of the validity of a short form of Wechsler's scales. *Psychol.Rep.* 26(2):559-561, Apr. 1970.

[32]Silverstein, A. B.: Reappraisal of the validity of WAIS, WISC, and WPPSI short forms. *J.Consult.&Clin.Psychol.* 34(1):12-14, Feb. 1970.

[33]Hite, L.: Analysis of Reliability and Validity of the Wechsler Intelligence Scale for Children. Unpublished doctoral dissertation. Western Reserve University, 1953.

[34]Hagen, E. P.: A Factor Analysis of the Wechsler Intelligence Scale for Children. Unpublished doctoral dissertation. Columbia University, 1952.

[35]Cohen, J.: The factorial structure of the WISC at ages 7-6, 10-6, and 13-6. *J.Consult.Psychol.* 23(4):285-299, Aug. 1959.

[36]Gault, U.: Factorial patterns of the Wechsler Intelligence Scales. *Aust.J.Psychol.* 6(1):85-89, June 1954.

[37]Seashore, H., Wesman, A., and Doppelt, J.: The standardization of the Wechsler Intelligence Scale for Children. *J.Consult.Psychol.* 14(2):99-110, Apr. 1950.

[38]Thorndike, R. L. (ed.): *Educational Measurement*, ed. 2. Washington. American Council on Education, 1971.

[39]Anastasi, A.: *Psychological Testing*, ed. 3. New York. The Macmillan Co., 1968.

[40]Anastasi, A.: *Individual Differences.* New York. John Wiley & Sons, Inc., 1965.

[41]Terman, L. M., and Merrill, M. A.: *Measuring Intelligence.* Boston. Houghton Mifflin Co., 1937.

[42]Schaie, K. W., and Strother, C. R.: A cross-sequential study of age changes in cognitive behavior. *Psychol.Bull.* 70(6):671-680, Dec. 1968.

[43]Tanner, J. M.: *Growth at Adolescence*, ed. 2. Oxford. Blackwell Scientific Publications, 1962.

[44]National Center for Health Statistics: Replication: An approach to the analysis of data from complex surveys. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 14. Public Health Service. Washington. U.S. Government Printing Office, Apr. 1966.

[45]National Center for Health Statistics: Pseudoreplication: Further evaluation and application of the balanced half-sample technique. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 31. Public Health Service. Washington. U.S. Government Printing Office, Jan. 1969.

LIST OF DETAILED TABLES

Page

Table 1. Mean and standard deviation (SD) of raw scores on the Vocabulary subtest of the WISC, by sex and age and mean scores, by grade, sex, and age for noninstitutionalized youths: United States, 1966-70 21

2. Mean, standard deviation (SD), coefficient of variation (CV), median, and standard error of the mean (SE) of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized youths, by 4-month age intervals: United States, 1966-70 22

3. Mean and standard deviation (SD) of raw scores on the Block Design subtest of the WISC, by sex and age, and mean scores, by grade, sex, and age for noninstitutionalized youths: United States, 1966-70 23

4. Mean scaled scores on the Vocabulary subtest of the WISC for noninstitutionalized youths, by grade, sex, and age: United States, 1966-70 24

5. Mean scaled scores on the Block Design subtest of the WISC for noninstitutionalized youths, by grade, sex, and age: United States, 1966-70 25

6. Scaled score equivalents for raw scores on the Vocabulary subtest of the WISC for noninstitutionalized youths, by 4-month age intervals: United States, 1966-70 26

7. Scaled score equivalents for raw scores on the Block Design subtest of the WISC for noninstitutionalized youths, by 4-month age intervals: United States, 1966-70 28

8. Conversion table for sum of scaled scores on the Vocabulary and Block Design subtests into estimates of Full Scale Scores among noninstitutionalized youths: United States, 1966-70 30

9. Mean estimated Full Scale Scores on the Vocabulary and Block Design short form of the WISC for noninstitutionalized youths, by grade, sex, and age: United States, 1966-70 31

10. Smoothed percentile equivalents of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized youths, by age and sex: United States, 1966-70 32

Table 1. Mean and standard deviation (SD) of raw scores on the Vocabulary subtest of the WISC, by sex and age, and mean scores, by grade, sex, and age, for noninstitutionalized youths: United States, 1966-70

Sex and age	Total		Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
	Mean	SD	Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
Both sexes			Mean raw score												
12-17 years	41.2	11.24	18.9	25.0	32.1	37.0	39.7	41.0	44.0	47.0	49.0	47.5	54.1	21.4	32.1
12 years	36.5	9.44	19.1	26.6	33.4	38.2	41.2	*	---	---	---	---	---	21.6	---
13 years	38.8	10.76	*	24.2	27.0	36.8	41.1	42.8	*	*	---	---	---	19.4	---
14 years	40.9	10.87	*	*	29.0	31.8	38.1	43.4	46.3	*	*	---	---	20.5	*
15 years	42.2	11.20	*	*	*	28.8	31.0	38.9	44.9	48.9	*	*	---	*	*
16 years	44.4	11.27	*	*	---	*	35.1	32.2	42.7	47.5	49.6	*	*	*	32.2
17 years	45.2	11.39	---	---	---	*	*	26.9	37.2	44.9	48.9	47.4	54.1	*	33.0
Boys															
12-17 years	42.2	11.26	20.5	25.0	33.8	38.5	41.2	42.2	45.1	48.0	50.2	50.1	54.6	22.2	33.6
12 years	37.8	9.75	20.5	25.4	35.0	39.9	43.2	*	---	---	---	---	---	22.4	---
13 years	40.3	10.96	*	25.0	29.6	38.7	43.0	44.1	*	---	---	---	---	20.0	---
14 years	42.0	11.01	*	*	29.6	33.2	39.7	45.5	48.6	---	*	---	---	21.1	*
15 years	43.5	11.12	*	---	*	31.3	32.8	40.3	47.1	51.0	*	*	---	*	*
16 years	45.1	11.34	---	---	---	*	37.2	33.6	42.7	48.8	52.0	---	*	*	30.8
17 years	45.2	11.49	---	---	---	*	*	26.8	37.5	45.1	49.8	50.1	54.9	*	35.8
Girls															
12-17 years	40.2	11.12	17.3	25.0	29.9	35.4	38.1	39.8	42.8	46.0	48.0	46.2	53.7	19.6	30.7
12 years	35.2	8.93	18.3	28.4	31.5	36.6	39.7	---	---	---	---	---	---	20.1	---
13 years	37.3	10.34	*	*	23.3	34.4	39.4	41.7	*	*	---	---	---	18.0	---
14 years	39.7	10.60	*	---	*	29.6	36.0	41.7	44.3	*	---	---	---	*	*
15 years	40.9	11.13	---	*	*	*	28.2	36.9	43.0	47.3	*	*	---	*	*
16 years	43.7	11.14	---	*	---	*	*	29.8	42.8	46.3	47.5	*	*	*	33.4
17 years	45.3	11.29	---	---	---	*	---	*	36.5	44.6	48.1	46.0	53.5	*	29.9
Standard error, total	0.59		3.82	2.05	1.01	0.66	0.54	0.63	0.64	0.43	0.53	1.91	1.27	1.21	0.84

Table 2. Mean, standard deviation (SD), coefficient of variation (CV), median, and standard error of the mean (SE) of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized youths, by 4-month age intervals: United States, 1966-70

Age interval	n	N	Vocabulary subtest					Block Design subtest				
			Mean	SD	CV	Median	SE	Mean	SD	CV	Median	SE
			Raw score									
12 years:												
0-3 months	318	1,081	35.6	8.98	0.25	36.8	0.55	23.8	12.88	0.54	23.5	0.62
4-7 months	440	1,466	36.1	9.47	0.26	36.4	0.75	23.5	12.66	0.66	23.1	0.94
8-11 months	432	1,456	37.6	9.65	0.26	38.0	0.40	24.8	13.04	0.53	25.1	0.70
13 years:												
0-3 months	414	1,356	38.6	10.98	0.28	39.0	0.65	26.0	13.02	0.50	27.3	0.75
4-7 months	403	1,323	38.2	10.61	0.28	38.5	0.84	25.9	13.75	0.53	26.5	1.01
8-11 months	390	1,271	39.8	10.61	0.27	40.8	0.87	27.2	13.51	0.50	27.9	0.68
14 years:												
0-3 months	414	1,337	40.4	10.64	0.26	41.3	0.86	28.6	13.51	0.47	29.2	0.63
4-7 months	379	1,170	40.9	11.12	0.27	40.5	1.32	29.3	13.90	0.47	29.8	0.99
8-11 months	409	1,334	41.3	10.87	0.26	42.2	0.69	29.9	13.53	0.45	30.6	0.74
15 years:												
0-3 months	370	1,220	41.8	11.12	0.27	42.1	0.67	29.0	13.03	0.45	29.4	0.66
4-7 months	381	1,287	42.3	11.33	0.27	42.4	0.80	30.1	14.01	0.47	30.5	0.87
8-11 months	365	1,244	42.5	11.15	0.26	42.8	1.11	29.5	13.62	0.46	30.8	0.83
16 years:												
0-3 months	404	1,408	43.9	10.55	0.24	45.0	0.89	31.5	13.66	0.43	33.3	0.95
4-7 months	328	1,029	44.4	11.97	0.27	45.6	0.82	32.1	13.74	0.43	33.5	0.76
8-11 months	360	1,188	44.9	11.44	0.25	45.4	0.80	31.9	14.64	0.46	34.5	0.94
17 years:												
0-3 months	336	1,232	45.0	11.63	0.26	46.0	0.93	33.0	13.86	0.42	34.9	0.83
4-7 months	294	1,061	44.8	10.93	0.24	46.2	0.83	32.5	13.67	0.42	34.4	0.82
8-11 months	327	1,214	45.9	11.51	0.25	46.4	0.98	33.4	13.70	0.41	35.7	0.72

NOTE.—n = sample size; N = estimated number of youths in population in thousands.

Table 3. Mean and standard deviation (SD) of raw scores on the Block Design subtest of the WISC, by sex and age, and mean scores, by grade, sex, and age for noninstitutionalized youths: United States, 1966-70

Sex and age	Total		Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
	Mean	SD	Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
			Mean raw score												
Both sexes															
12-17 years	28.9	13.88	10.3	13.0	20.3	24.6	27.4	29.4	31.4	34.2	35.1	37.1	40.2	12.0	22.2
12 years	24.0	12.87	9.2	14.5	21.0	25.7	28.6	*	---	---	---	---	---	9.6	---
13 years	26.3	13.44	*	10.9	17.5	24.4	28.0	31.0	*	*	---	---	---	11.8	---
14 years	29.3	13.65	*	*	20.4	19.2	27.1	31.4	34.5	*	*	---	---	10.8	*
15 years	29.6	13.58	*	*	*	18.7	20.9	27.1	31.9	33.2	*	*	---	*	*
16 years	31.8	14.01	*	*	---	*	23.7	23.2	29.4	35.0	35.4	*	*	*	21.4
17 years	33.0	13.75	---	---	---	*	*	19.8	29.7	33.0	35.0	37.2	39.9	*	22.8
Boys															
12-17 years	30.4	13.73	14.5	13.5	22.3	25.8	29.3	31.0	33.7	36.2	37.3	38.1	43.4	13.8	24.7
12 years	25.6	12.82	13.7	14.8	22.8	27.2	33.0	*	---	---	---	---	---	11.0	---
13 years	27.7	13.32	*	10.5	20.4	26.0	29.6	32.1	*	---	---	---	---	14.7	---
14 years	30.8	13.86	*	*	24.4	19.6	29.4	33.8	36.7	---	*	---	---	*	*
15 years	31.3	13.21	*	---	*	20.0	22.6	29.4	34.6	33.6	*	*	---	*	*
16 years	34.2	13.44	---	---	---	*	24.8	24.5	31.7	37.8	39.4	---	*	*	22.7
17 years	33.9	13.64	---	---	---	*	*	19.6	30.8	34.1	36.8	37.7	42.8	*	26.2
Girls															
12-17 years	27.2	13.84	6.1	12.3	17.8	23.3	25.4	27.7	29.2	32.3	33.1	36.6	37.5	8.3	19.8
12 years	22.4	12.71	6.8	14.0	18.9	24.2	25.2	---	---	---	---	---	---	6.7	---
13 years	25.0	13.42	*	*	13.6	22.4	26.6	30.0	*	*	---	---	---	6.2	---
14 years	27.6	13.24	*	---	*	18.5	24.1	29.4	32.6	*	---	---	---	*	*
15 years	27.7	13.70	---	*	*	*	18.3	23.9	29.5	32.8	*	*	---	*	*
16 years	29.4	14.17	*	*	---	*	*	20.8	27.0	32.3	31.8	*	*	*	20.2
17 years	32.0	13.80	---	---	---	*	---	*	27.3	31.9	33.4	37.0	37.5	*	18.9
Standard error, total	0.36		2.84	1.72	1.08	0.63	0.40	0.42	0.48	0.52	0.46	1.59	1.61	0.91	1.24

Table 4. Mean Scaled Scores on the Vocabulary subtest of the WISC for noninstitutionalized youths, by grade, sex, and age: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
	Mean Scaled Score													
Both sexes														
12-17 years	10.0	4.5	6.5	8.7	9.9	10.1	9.9	10.3	10.7	10.9	10.5	12.4	5.0	6.8
12 years	10.0	4.9	7.2	9.2	10.6	11.2	*	---	---	---	---	---	5.7	---
13 years	10.0	*	6.0	6.8	9.5	10.6	10.9	*	*	---	---	---	4.6	---
14 years	10.0	*	*	6.7	7.5	9.3	10.7	11.3	*	*	---	---	4.5	*
15 years	10.0	*	*	*	6.5	7.0	9.1	10.7	11.8	*	*	---	*	*
16 years	10.0	*	*	---	*	*	6.9	9.6	10.8	11.3	*	*	*	6.9
17 years	10.0	---	---	---	*	*	5.6	8.0	9.9	10.8	10.4	12.3	*	6.9
Boys														
12-17 years	10.3	4.8	6.6	9.1	10.3	10.5	10.2	10.6	11.0	11.3	11.2	12.6	5.2	7.1
12 years	10.4	5.5	6.8	9.7	11.0	11.8	*	---	---	---	---	---	5.9	---
13 years	10.4	*	6.2	7.6	10.0	11.2	11.2	*	---	---	---	---	4.8	---
14 years	10.3	*	*	6.9	7.9	9.7	11.2	12.0	---	*	---	---	4.6	*
15 years	10.4	---	---	*	7.2	7.5	9.5	11.3	12.4	*	*	---	*	*
16 years	10.2	---	---	---	*	*	7.3	9.6	11.2	12.0	---	*	*	6.5
17 years	10.0	---	---	---	*	*	5.6	8.1	9.9	11.1	11.2	12.6	*	7.6
Girls														
12-17 years	9.7	4.2	6.4	8.0	9.5	9.7	9.6	10.0	10.4	10.6	10.0	12.2	4.6	6.5
12 years	9.6	4.6	7.6	8.6	10.1	10.7	- -	---	---	---	---	---	5.4	---
13 years	9.6	*	*	5.9	8.8	10.1	10.6	*	*	---	---	---	4.2	---
14 years	9.6	*	---	*	6.9	8.7	10.2	10.8	*	---	---	---	*	*
15 years	9.6	---	*	*	*	6.2	8.6	10.2	11.3	*	*	---	*	*
16 years	9.8	---	*	---	*	*	6.2	9.6	10.4	10.8	*	*	*	7.3
17 years	10.0	---	---	---	*	---	*	7.9	9.9	10.6	10.0	12.1	*	6.1
Standard error, total	0.16	1.18	0.62	0.29	0.19	0.14	0.17	0.16	0.13	0.14	0.53	0.33	0.37	0.21

Table 5. Mean Scaled Scores on the Block Design subtest of the WISC for noninstitutionalized youths, by grade, sex, and age: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
	Mean Scaled Score													
Both sexes														
12-17 years	10.0	5.8	7.3	9.1	9.9	10.1	10.0	10.2	10.5	10.5	10.8	11.5	6.2	8.0
12 years	10.0	5.8	7.8	9.4	10.4	10.9	•	---	---	---	---	---	6.6	---
13 years	10.0	•	6.6	8.2	9.6	10.3	10.8	•	•	---	---	---	5.9	---
14 years	10.0	•	•	8.2	7.9	9.6	10.5	11.1	•	•	---	---	5.1	•
15 years	10.0	•	•	•	7.4	8.0	9.4	10.4	10.7	•	•	---	•	•
16 years	10.0	•	•	---	•	8.2	8.3	9.5	10.6	10.7	•	•	•	7.8
17 years	10.0	---	---	---	•	•	7.5	9.4	10.0	10.5	10.9	11.4	•	8.0
Boys														
12-17 years	10.4	7.0	7.5	9.6	10.1	10.4	10.3	10.7	11.0	11.0	11.0	12.2	6.6	8.5
12 years	10.4	7.3	8.0	9.8	10.7	11.9	•	---	---	---	---	---	7.0	---
13 years	10.2	•	6.6	8.8	9.9	10.7	11.2	•	---	---	---	---	6.6	---
14 years	10.4	•	•	9.0	8.0	10.1	11.0	11.6	---	•	---	---	•	•
15 years	10.3	•	---	•	7.9	8.5	9.8	11.0	10.8	•	•	---	•	•
16 years	10.5	---	---	---	•	8.4	8.5	10.0	11.3	11.6	---	•	•	8.2
17 years	10.3	---	---	---	•	•	7.5	9.7	10.3	10.8	11.0	11.9	•	8.7
Girls														
12-17 years	9.6	4.6	6.9	8.6	9.6	9.7	9.7	9.7	10.0	10.1	10.7	11.0	5.3	7.5
12 years	9.6	5.1	7.6	8.9	10.1	10.1	---	---	---	---	---	---	5.6	---
13 years	9.6	•	•	7.3	9.1	10.0	10.5	•	•	---	---	---	4.6	---
14 years	9.7	•	---	•	7.8	9.0	10.1	10.6	•	---	---	---	•	•
15 years	9.5	---	•	•	•	7.3	8.7	9.9	10.6	•	•	---	•	•
16 years	9.4	•	•	---	•	•	7.9	9.0	10.0	9.9	•	•	•	7.5
17 years	9.8	---	---	---	•	---	•	8.8	9.8	10.1	10.8	11.0	•	7.3
Standard error, total	0.08	1.06	0.48	0.25	0.14	0.09	0.09	0.10	0.12	0.11	0.31	0.41	0.24	0.32

Table 6. Scaled Score equivalents for raw scores on the Vocabulary subtest of the WISC for noninstutionalized youths, by 4-month age intervals: United States, 1966-70

Scaled Scores	12 years			13 years			14 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Raw scores								
0	0-5	0-6	0-6	0-7	0-7	0-7	0-8	0-8	0-8
1	6-7	7-8	7-8	8-9	8-9	8-9	9-10	9-10	9-10
2	8-10	9-11	9-11	10-11	10-12	10-12	11-13	11-13	11-13
3	11-14	12-14	12-14	12-14	13-14	13-14	14-15	14-15	14-15
4	15-18	15-18	15-18	15-18	15-18	15-18	16-18	16-19	16-19
5	19-20	19-21	19-21	19-21	19-21	19-22	19-22	20-22	20-24
6	21-24	22-24	22-24	22-24	22-25	23-27	23-27	23-28	25-28
7	25-27	25-27	25-28	25-28	26-29	28-31	28-32	29-32	29-32
8	28-31	28-31	29-33	29-33	30-33	32-35	33-35	33-35	33-35
9	32-34	32-34	34-36	34-36	34-37	36-39	36-39	36-39	36-40
10	35-38	35-38	37-39	37-40	38-40	40-42	40-42	40-42	41-43
11	39-40	39-41	40-41	41-43	41-43	43-45	43-45	43-46	44-46
12	41-42	42-44	42-45	44-48	44-48	46-48	46-48	47-51	47-51
13	43-45	45-46	46-48	49-50	49-50	49-51	49-52	52-54	52-54
14	46-48	47-49	49-52	51-53	51-53	52-54	53-56	55-57	55-57
15	49-50	50-52	53-55	54-56	54-56	55-57	57-59	58-59	58-60
16	51-54	53-56	56-58	57-59	57-60	58-60	60-61	60-62	61-63
17	55-56	57-59	59-60	60-62	61-63	61-63	62-64	63-65	64-65
18	57-59	60-61	61-62	63-64	64-65	64-65	65-66	66-67	66-67
19	60-61	62-63	63-64	65-66	66-67	66-67	67-68	68-69	68-69
20	62-80	64-80	65-80	67-80	68-80	68-80	69-80	70-80	70-80

Table 6. Scaled Score equivalents for raw scores on the Vocabulary subtest of the WISC for noninstitutionalized youths, by 4-month age intervals: United States, 1966-70—Con.

Scaled Scores	15 years			16 years			17 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Raw scores								
0	0-9	0-9	0-9	0-10	0-10	0-10	0-10	0-10	0-10
1	10-11	10-12	10-12	11-12	11-12	11-12	11-12	11-12	11-12
2	12-13	13-15	13-15	13-15	13-15	13-15	13-15	13-15	13-15
3	14-16	16-18	16-18	16-18	16-18	16-18	16-18	16-18	16-19
4	17-21	19-21	19-21	19-21	19-21	19-21	19-21	19-21	20-22
5	22-24	22-24	22-25	22-26	22-26	22-26	22-26	22-26	23-26
6	25-28	25-28	26-29	27-30	27-30	27-31	27-31	27-31	27-31
7	29-32	29-32	30-32	31-34	31-34	32-34	32-34	32-35	32-35
8	33-35	33-36	33-36	35-38	35-38	35-38	35-38	36-40	36-41
9	36-40	37-40	37-40	39-43	39-43	39-43	39-43	41-44	42-45
10	41-43	41-44	41-44	44-46	44-47	44-47	44-48	45-48	46-48
11	44-48	45-48	45-48	47-49	48-51	48-51	49-51	49-51	49-51
12	49-52	49-52	49-52	50-52	52-55	52-55	52-55	52-55	52-55
13	53-55	53-56	53-56	53-57	56-58	56-58	56-58	56-58	56-59
14	56-58	57-58	57-58	58	59-60	59-60	59-60	59-60	60-61
15	59-61	59-61	59-61	59-61	61-63	61-63	61-63	61-63	62-64
16	62-63	62-63	62-63	62-64	64	64-65	64-65	64-66	65-66
17	64-66	64-66	64-66	65-66	65-66	66-67	66-67	67-68	67-68
18	67-68	67-68	67-68	67-68	67-68	68-69	68-69	69-70	69-70
19	69	69-70	69-70	69-70	69-70	70	70-71	71	71
20	70-80	71-80	71-80	71-80	71-80	71-80	72-80	72-80	72-80

Table 7. Scaled Score equivalents for raw scores on the Block Design subtest of the WISC for noninstitutionalized youths, by 4-month age intervals: United States, 1966-70

Scaled Scores	12 years			13 years			14 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Raw scores								
0	-	-	-	-	-	-	-	-	-
1	-	-	0-1	0-1	0-1	0-1	0-1	0-1	0-1
2	0-1	0-1	2-3	2-3	2-3	2-3	2-3	2-3	2-3
3	2-3	2-3	4	4	4	4	4	4	4
4	4	4	5	5	5	5	5	5	5
5	5	5	6	6	6	6-7	6-7	6-8	6-8
6	6-8	6-9	7-9	7-9	7-9	8-10	8-10	9-10	9-11
7	9-10	10	10-11	10-11	10-11	11	11-13	11-13	12-14
8	11-13	11-13	12-14	12-18	12-18	12-19	14-21	14-21	15-22
9	14-20	14-21	15-22	19-22	19-23	20-25	22-26	22-26	23-27
10	21-25	22-25	23-27	23-29	24-29	26-30	27-31	27-32	28-33
11	26-31	26-31	28-32	30-34	30-35	31-35	32-37	33-38	34-38
12	32-37	32-37	33-37	35-38	36-39	36-40	38-41	39-43	39-43
13	38-40	38-40	38-41	39-42	40-44	41-44	42-45	44-46	44-46
14	41-42	41-43	42-45	43-45	45-47	45-47	46-47	47-48	47-49
15	43-45	44-45	46-47	46-47	48-49	48-50	48-50	49-51	50-51
16	46-49	46-49	48-50	48-50	50	51-52	51-52	52	52
17	50	50-51	51-52	51-52	51-52	53	53	53	53
18	51-52	52	53	53	53	54	54	54	54
19	53	53	54	54	54	55	55	55	55
20	54-55	54-55	55	55	55	-	-	-	-

Table 7. Scaled Score equivalents for raw scores on the Block Design subtest of the WISC for noninstitutionalized youths, by 4-month age intervals: United States, 1966-70—Con.

Scaled Scores	15 years			16 years			17 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Raw scores								
0	-		0	0	0	0	0	0	0-1
1	0-1	0-1	1	1	1	1	1	1-2	2
2	2-3	2-3	2-3	2-3	2-3	2-3	2-3	3	3
3	4	4	4	4	4	4	4	4	4
4	5	5	5	5	5	5	5	5	5
5	6-9	6-9	6-9	6-9	6-9	6-9	6-9	6-9	6-9
6	10-11	10-11	10-11	10-11	10-11	10-11	10-12	10-12	10-12
7	12-14	12-14	12-15	12-18	12-19	12-19	13-19	13-20	13-20
8	15-22	15-22	16-22	19-24	20-26	20-26	20-26	21-26	21-27
9	23-27	23-27	23-28	25-30	27-31	27-31	27-31	27-31	28-33
10	28-33	28-33	29-33	31-36	32-36	32-37	32-37	32-37	34-38
11	34-38	34-38	34-38	37-40	37-41	38-41	38-42	38-42	39-42
12	39-43	39-43	39-43	41-44	42-44	42-45	43-45	43-45	43-46
13	44-46	44-47	44-47	45-47	45-47	46-48	46-48	46-48	47-48
14	47-49	48-50	48-50	48-50	48-50	49-50	49-50	49-50	49-50
15	50-51	51	51	51	51	51-52	51-52	51-52	51-52
16	52	52	52	52	52	53	53	53	53
17	53	53	53	53	53	54	54	54	54
18	54	54	54	54-55	54-55	55	55	55	55
19	55	55	55	-	-	-	-	-	-
20	-	-	-	-	-	-	-	-	-

Table 8. Conversion table for sum of Scaled Scores on the Vocabulary and Block Design subtests into estimates of Full Scale Scores among noninstitutionalized youths: United States, 1966-70

Sum of Scaled Scores	Short-form estimate of Full Scale Score	Sum of Scaled Scores	Short-form estimate of Full Scale Score
00	43	21	103
01	46	22	106
02	49	23	109
03	51	24	111
04	54	25	114
05	57	26	117
06	60	27	120
07	63	28	123
08	66	29	126
09	69	30	129
10	71	31	131
11	74	32	134
12	77	33	137
13	80	34	140
14	83	35	143
15	86	36	146
16	89	37	149
17	91	38	151
18	94	39	154
19	97	40	157
20	100		

Table 9. Mean estimated Full Scale Scores on the Vocabulary and Block Design short form of the WISC for noninstitutionalized youths, by grade, sex, and age: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
	Mean estimated Full Scale Score													
Both sexes														
12-17 years	100.0	72.4	82.3	93.7	99.4	100.5	99.8	101.4	103.4	104.2	103.7	111.2	74.9	85.1
12 years	100.1	73.5	85.7	96.1	102.7	106.0	*	...	...	...	...	...	78.0	...
13 years	99.9	*	78.9	85.8	97.4	102.8	104.9	*	*	...	...	...	73.1	...
14 years	100.1	*	*	85.2	87.0	97.0	103.3	106.9	*	*	...	...	70.0	*
15 years	99.7	*	*	*	82.8	85.7	95.8	103.2	107.1	*	*	...	*	*
16 years	100.0	*	*	...	*	88.6	86.3	97.6	104.2	105.8	*	*	*	85.0
17 years	100.0	...	...	...	*	*	80.2	92.4	99.8	103.7	103.7	110.7	*	85.5
Boys														
12-17 years	101.9	76.9	83.3	96.4	101.2	102.7	101.5	103.7	105.5	106.5	106.5	113.9	76.6	87.5
12 years	102.3	79.6	85.2	98.6	105.1	110.6	*	...	...	...	...	...	79.9	...
13 years	102.0	*	79.3	89.6	99.9	105.4	106.7	*	...	...	...	...	75.7	...
14 years	101.9	*	*	88.0	88.4	99.5	106.3	110.3	...	*	...	...	70.5	*
15 years	101.9	...	...	*	86.0	88.5	98.2	106.5	109.0	*	*	...	*	*
16 years	102.2	...	...	...	*	90.7	88.0	99.0	107.1	110.2	...	*	*	84.9
17 years	100.7	...	...	...	*	*	80.4	93.4	100.6	105.5	106.2	113.2	*	89.3
Girls														
12-17 years	98.0	67.9	81.0	90.3	97.4	98.2	98.1	99.2	101.5	102.1	102.4	108.9	71.3	82.8
12 years	97.8	70.2	86.4	92.9	100.4	102.5	...	...	...	...	...	...	74.5	...
13 years	97.7	*	*	80.4	94.2	100.3	103.4	*	*	...	...	...	67.8	...
14 years	98.1	*	...	*	84.8	93.5	100.8	104.0	*	...	...	...	*	*
15 years	97.5	...	*	*	*	81.5	92.4	100.2	105.6	*	*	...	*	*
16 years	97.9	...	*	...	*	*	83.3	96.2	101.3	101.9	*	*	*	85.1
17 years	99.4	...	...	...	*	...	*	90.3	99.1	102.1	102.4	108.6	*	81.1
Standard error, total	0.65	6.00	3.00	1.47	0.91	0.55	0.71	0.66	0.63	0.62	2.17	1.88	1.76	1.17

Table 10. Smoothed percentile equivalents of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized youths, by age and sex: United States, 1966-70

Sex and percentile	Age in years											
	12	13	14	15	16	17	12	13	14	15	16	17
Both sexes	Vocabulary raw score						Block Design raw score					
99	58	61	64	65	66	67	51	51	53	53	54	54
98	55	59	61	63	64	66	49	50	52	52	53	53
97	54	57	60	62	63	65	47	49	51	52	53	53
96	53	56	59	61	62	64	46	48	50	51	52	52
95	52	55	58	60	61	63	45	47	50	50	52	52
90	48	53	55	57	59	59	42	45	47	47	49	50
85	46	50	52	54	56	57	39	42	45	45	47	48
80	44	48	50	53	54	55	37	40	43	43	45	46
75	43	47	48	51	53	53	34	37	41	41	43	44
70	42	45	47	49	51	52	32	35	39	39	41	43
65	41	43	45	47	49	50	30	33	37	37	40	41
60	40	42	44	45	48	49	28	31	34	35	38	39
55	38	41	43	44	47	48	26	29	32	33	36	37
50	37	39	41	42	45	46	24	27	30	30	34	35
45	36	38	40	41	44	45	22	25	28	28	32	33
40	34	37	38	39	42	43	20	23	26	26	30	31
35	33	35	37	38	40	42	17	20	24	24	28	29
30	32	33	36	36	39	40	13	17	22	22	25	26
25	30	32	34	34	37	37	12	13	18	20	22	23
20	28	30	32	33	35	35	11	12	14	14	18	20
15	26	27	30	30	32	33	10	10	11	12	12	15
10	24	24	27	27	29	30	6	9	10	10	10	11
5	20	20	22	24	25	25	6	6	6	6	6	6
4	19	19	20	23	24	24	5	6	6	6	6	6
3	18	18	19	21	22	22	5	5	6	6	6	6
2	16	17	17	19	20	20	4	5	5	5	5	6
1	14	14	14	16	17	17	4	4	4	4	4	5
Boys												
99	59	62	64	64	66	67	52	52	54	54	54	54
98	57	60	62	63	65	66	49	50	52	52	54	54
97	56	59	60	62	64	65	48	49	52	52	53	53
96	55	58	60	61	63	64	47	49	51	52	53	53
95	54	57	59	60	62	63	46	48	50	51	53	53
90	50	54	56	58	59	59	43	46	47	49	50	51
85	48	52	54	56	58	58	41	43	46	46	49	49
80	46	50	52	54	56	56	39	41	44	44	47	47
75	44	49	50	53	54	54	36	38	43	43	45	45
70	43	47	48	50	52	52	33	37	41	41	43	43
65	42	45	47	49	51	51	32	34	40	40	42	42
60	41	43	45	47	49	49	30	32	38	38	40	41
55	40	42	44	45	47	48	28	30	35	35	39	39
50	38	41	42	44	46	46	26	29	33	33	36	37
45	37	39	41	42	44	44	24	27	30	30	34	35
40	36	38	39	41	43	43	22	25	28	29	32	32
35	34	36	38	39	41	41	20	22	25	27	30	30
30	33	34	37	37	39	40	17	19	23	25	28	28
25	31	33	35	35	38	38	13	14	20	22	25	25
20	29	31	34	34	35	35	12	12	16	19	22	22

Table 10. Smoothed percentile equivalents of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized youths, by age and sex: United States, 1966-70—Con.

Sex and percentile	Age in years											
	12	13	14	15	16	17	12	13	14	15	16	17
Boys—Con.	Vocabulary raw score						Block Design raw score					
15	27	29	32	32	33	33	11	12	12	13	18	18
10	25	26	28	29	29	30	10	10	10	12	12	12
5	21	21	22	25	25	26	06	06	06	10	10	10
4	20	20	20	24	24	24	06	06	06	06	06	06
3	19	19	19	23	23	23	06	06	06	06	06	06
2	17	18	18	21	21	21	05	05	06	06	06	06
1	14	14	14	17	17	17	04	04	04	04	05	06
Girls												
99	54	58	63	66	66	67	50	50	53	53	53	54
98	53	56	60	63	63	66	47	49	51	52	53	53
97	51	55	58	61	62	64	46	48	50	50	52	52
96	50	54	57	60	61	63	45	48	50	50	51	52
95	49	53	56	58	61	62	44	47	49	49	50	52
90	46	50	53	55	57	59	40	44	46	46	46	49
85	44	48	50	53	55	57	38	41	43	43	44	47
80	43	46	48	51	53	55	34	38	41	41	43	46
75	42	45	47	49	51	53	32	36	38	40	41	44
70	41	43	46	47	50	52	30	34	36	38	39	41
65	39	42	44	45	48	51	28	31	34	35	37	39
60	38	41	43	44	47	49	26	29	32	33	36	37
55	37	39	42	43	46	48	24	27	30	30	33	35
50	36	38	41	41	45	46	22	25	28	28	31	34
45	35	37	39	39	43	45	21	23	26	26	29	32
40	33	36	37	38	42	44	17	21	25	25	27	30
35	32	34	36	36	40	42	13	18	22	22	24	28
30	31	33	34	35	39	41	12	13	20	20	22	25
25	29	31	33	33	36	38	11	12	16	16	18	23
20	28	28	31	32	34	35	10	11	12	13	13	18
15	25	26	28	29	32	33	09	10	11	11	11	13
10	23	23	25	25	28	30	06	06	10	10	10	1[illegible]
5	19	19	21	23	24	24	05	06	06	06	06	06
4	18	19	20	21	23	23	05	05	05	05	05	06
3	17	17	18	19	20	22	05	05	05	05	05	06
2	16	16	16	18	19	19	04	05	05	05	05	06
1	14	14	14	14	16	16	03	04	04	04	04	05

APPENDIX I
TECHNICAL NOTES

The Survey Design

The sample designs for the first three programs, or Cycles I-III, of the Health Examination Survey have been essentially similar, in that each has been a multistage, stratified probability sample of clusters of households in land-based segments. The successive elements for this sample design are primary sampling unit (PSU), census enumeration district (ED), segment (a cluster of households), household, eligible youths, and finally, the sample youth.

The 40 sample areas and the segments utilized in the design of Cycle III were the same as those in Cycle II. Previous reports describe in detail the sample design used for Cycle II and in addition discuss the problems and considerations given to other types of sampling frames and whether or not to control the selection of siblings.[6,7]

Requirements and limitations placed on the design for Cycle III, similar to those for the design for Cycle II, were that:

1. The target population be defined as the civilian noninstitutional population of the United States, including Alaska and Hawaii, in the age range of 12 through 17 years, with the special exclusion of children residing on reservation lands of the American Indians. The latter exclusion was adopted as a result of operational problems encountered on these lands in Cycle I.
2. The time period of data collection be limited to about 3 years for each cycle and the length of the individual examination within the specially constructed mobile examination center be between 2 and 3 hours.
3. Ancillary data be collected on specially designed household, medical history, and school questionnaires, and from copies of birth certificates.
4. Examination objectives be related primarily to factors of physical and intellectual growth and development.
5. The sample be sufficiently large to yield reliable findings within broad geographic regions and population density groups as well as within age, sex, and limited socioeconomic groups for the total sample.

The sample was drawn jointly with the U.S. Bureau of the Census, beginning with the 1960 Decennial Census list of addresses and the nearly 1,900 PSU's into which the entire United States was divided. Each PSU is either a standard metropolitan statistical area (SMSA), a county, or a group of two or three contiguous counties. These PSU's were grouped into 40 strata, with each stratum having an average size of about 4.5 million persons. Stratification was accomplished so as to maximize the degree of homogeneity within strata with regard to the population size of the PSU's, degree of urbanization, geographic proximity, and degree of industrialization. The 40 strata were then classified into four broad geographic regions of 10 strata each and then within each region, cross-classified by four population density classes and classes of rate of population change from 1950 to 1960. Using a modified Goodman-Kish controlled-selection technique, one PSU was drawn from each of the 40 strata.

Generally, within each PSU, 20 census enumeration districts were selected, with the probability of selection of a particular ED proportional to its population in the age group 5-9 years in the 1960 Census, which by 1966 approximated the target population for Cycle III. A similar method was used for selecting one segment (a smaller cluster of households) in each

ED. Because of the approximately 3-year time interval between Cycle II and Cycle III, the Cycle III sampling frame was updated for new construction and to compensate for segments where housing was partially or totally demolished to make room for highway construction or urban redevelopment. Each of the resulting 20 segments within a PSU was either a bounded area or a cluster of households (or addresses). All youths in the appropriate age range who resided at the address visited were eligible youths, i.e., eligible for inclusion in the sample. Operational considerations made it necessary to reduce the number of prospective examinees at any one location to a maximum of 200. When the number of eligible youths in a particular location exceeded this number, the "excess" eligible youths were deleted from the sample through a systematic sampling technique. Youths who were not selected as sample youths in the Cycle III sample, but who were previously examined in Cycle II, were scheduled for examination if time permitted and will be included in special longitudinal analyses. In addition, individual twins who were deleted from the Cycle III sample were also scheduled for examination, as they were in Cycle II, to provide data on pairs of twins for future analysis. These data are not included in the report as part of the national probability sample of youths.

The sample was selected in Cycle III, as it had been for the children in Cycle II, to contain proportional representation of youths from families having only one eligible youth, two eligible youths, and so on, so as to be representative of the total target population. However, since households were one of the elements in the sample frame, the number of related youths in the resulting sample is greater than would result from a design which sampled youths 12-17 years without regard to household. The resulting estimated mean measurements or rates should be unbiased but their sampling variabilities will be somewhat greater than those from a more costly, time-consuming, systematic sample design in which every kth youth would be selected.

The total probability sample for Cycle III included 7,514 youths representative of the approximately 22.7 million noninstitutionalized United States youths of 12-17 years. The sample contained youths from 25 different States, with approximately 1,000 in each single year of age.

The response rate in Cycle III was 90 percent, with 6,768 youths examined out of the total sample. These examinees were closely representative of those in the population from which the sample was drawn with respect to age, sex, race, region, population density, and population growth in area of residence. Hence it appears unlikely that nonresponse could bias the findings appreciably.

Reliability

While measurement processes in the surveys were carefully standardized and closely controlled, the correspondence between true population figures and HES results cannot be expected to be exact. Survey data are imperfect for three major reasons: (1) results are subject to sampling error, (2) the actual conduct of a survey never agrees perfectly with the design, and (3) the measurement processes themselves are inexact, even though standardized and controlled.

Data recorded for each sample youth are inflated in the estimation process to characterize the larger universe of which the sample youths are representative. The weights used in this inflation process are a product of the reciprocal of the probability of selecting the youth, an adjustment for nonresponse cases, and a poststratified ratio adjustment that increases precision by bringing survey results into closer alinement with known U.S. population figures by color and sex within single years of age for ages 12-17.

In the third cycle of the Health Examination Survey, as for the children in Cycle II, the sample was the result of three principal stages of selection: the single PSU from each stratum, the 20 segments from each sample PSU, and the sample youth from the eligible youths. The probability of selecting an individual youth is the product of the probability of selection at each stage.

Because the strata are roughly equal in population size and a nearly equal number of sample youths were examined in each of the sample PSU's, the sample design is essentially self-weighting with respect to the target population; that is, each youth 12 through 17 years of age had about the same probability of being drawn into the sample.

The adjustment upward for nonresponse is intended to minimize the impact of nonresponse on final estimates by imputing to nonrespondents the characteristics of "similar" respondents. Here, "similar" respondents in a sample PSU were defined as examined youths of the same age in years and sex as youths not examined in that sample PSU.

The poststratified ratio adjustment used in the third cycle achieved most of the gains in precision that would have been attained if the sample had been drawn from a population stratified by age, color, and sex and makes the final sample estimates of population agree exactly with independent controls prepared by the Bureau of the Census for the U.S. noninstitutional population as of March 9, 1968, approximate midpoint of the survey for Cycle III, by color and sex for each single year of ages 12-17. The weight of every responding sample youth in each of the 24 age, color, and sex classes is adjusted upward or downward so that the weighted total within the class equals the independent population control. Final sample frequencies and estimated population frequencies as of the approximate midpoint of the survey are presented in table I by age and sex.

Extent of Missing Test Results and Imputation Procedures Used

In addition to youths who were selected for the sample but, for various reasons, not examined, there were some whose examination was incomplete in one procedure or another. The extent of missing data for the WISC is shown in table II according to sex and age. For 109 youths, or 1.6 percent of all those examined, one or both of the WISC subtest results were not available. There were a number of reasons for this, the primary one being operational and logistical survey problems, such as lost records or lack of time to complete the examination. Other reasons included difficulties with the English language or illness on the part of the youth. Since the reason for missing test results in most cases was not directly related to the characteristic being measured, raw scores were imputed for almost all of these examinees. In certain, infrequent instances, imputation was not considered appropriate, as for example, the imputation of Block Design scores for a blind youth, or the imputation of Vocabulary test results for foreign-language-speaking youths who could not understand English well enough to take any of the psychological tests.

Imputation was accomplished in the following manner. An intercorrelation matrix of all psychological test data and selected socioeconomic variables was derived to identify those variables that were most highly associated with each raw test score. As a result, five variables were chosen for the imputation of Vocabulary and Block Design raw scores: other available test scores, educational level of the head of the household (four categories), age, and two con-

Table I. Sample and estimated population frequency distributions of youths 12-17 years of age in the noninstitutional population of the United States: Health Examination Survey, 1966-70

Age	Number of youths in sample			Estimated number of youths in population as of midsurvey		
	Total	Boys	Girls	Total	Boys	Girls
				Number in thousands		
Total	6,768	3,545	3,223	22,692	11,489	11,203
12 years	1,190	643	547	4,002	2,032	1,970
13 years	1,208	626	582	3,952	2,006	1,946
14 years	1,204	618	586	3,852	1,951	1,901
15 years	1,116	613	503	3,751	1,900	1,851
16 years	1,092	556	536	3,625	1,836	1,789
17 years	958	489	469	3,510	1,764	1,746

Table II. Extent of missing WISC test results among examinees in the Health Examination Survey, 1966-70

Sex and age	Total	Both subtests missing	Vocabulary only missing	Block Design only missing
	Number of examinees			
Both sexes				
12-17 years	109	71	17	21
12 years	18	12	3	3
13 years	16	8	3	5
14 years	24	15	4	5
15 years	15	11	3	1
16 years	18	11	3	4
17 years	18	14	1	3
Boys				
12-17 years	60	40	7	13
12 years	9	7	1	1
13 years	10	5	1	4
14 years	17	12	2	3
15 years	8	8	-	-
16 years	7	3	2	2
17 years	9	5	1	3
Girls				
12-17 years	49	31	10	8
12 years	9	5	2	2
13 years	6	3	2	1
14 years	7	3	2	2
15 years	7	3	3	1
16 years	11	8	1	2
17 years	9	9	-	-

trol variables, race and sex. Imputation of a missing test result for an examinee was accomplished by randomly selecting a match among the group of examinees of the same age in years, parental level of education (four categories), race, sex, and available raw score test results most highly correlated with the scores to be imputed. The raw score of this "matched" examinee was then imputed to the examinee with the missing score. When data for any of these variables were not available, a match was selected using information on as many of the variables as were available in the youth's record.

Sampling and Measurement Error

In the present report, reference has been made to efforts to minimize bias and variability of measurement techniques. The probability design of the survey makes possible the calculation of sampling errors. The sampling error is used here to determine how imprecise the survey test results may be because they result from a sample rather than from the measurements of all elements in the universe. The estimation of sampling errors for a study of the type of the Health Examination Survey is difficult for at

least three reasons: (1) measurement error and "pure" sampling error are confounded in the data, and it is difficult to find a procedure that will either completely include both or treat one or the other separately; (2) the survey design and estimation procedure are complex, and accordingly, require computationally involved techniques for the calculation of variances; and (3) thousands of statistics are derived from the survey, many for subclasses of the population for which the number of sample cases is small. Estimates of sampling error are obtained from the sample data and are themselves subject to sampling error, which may be large when the number of cases in a cell is small or, occasionally, even when the number of cases is substantial.

Estimates of approximate sampling variability for selected statistics used in this report are included in the detailed tables and in tables III through VII. These estimates, called standard errors, have been prepared by a replication technique that yields overall variability through observation of variability among random subsamples of the total sample. The method reflects both "pure" sampling variance and a part of the measurement variance, and is described in previously published reports.[44,45]

Hypothesis Testing

In accordance with usual practice, the interval estimate for any statistic was considered to be

Table III. Standard errors for mean raw scores on the Vocabulary subtest of the WISC by grade, sex, and age for noninstitutionalized youths: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
Both sexes	Standard error													
12-17 years	0.59	3.82	2.05	1.01	0.66	0.54	0.63	0.64	0.43	0.53	1.91	1.27	1.21	0.84
12 years	0.50	5.08	2.00	0.94	0.36	0.92	*	---	---	---	---	---	1.76	---
13 years	0.62	*	4.07	1.82	1.58	0.45	1.00	*	*	---	---	---	2.77	---
14 years	0.81	*	*	2.86	1.32	1.32	0.61	1.09	*	*	---	---	2.76	*
15 years	0.67	*	*	*	2.77	2.53	1.05	0.64	1.34	*	*	---	*	*
16 years	0.57	*	*	---	*	2.16	1.22	1.38	0.33	1.18	*	*	*	1.60
17 years	0.69	---	---	---	*	*	2.36	1.18	1.06	0.51	1.98	1.44	*	0.65
Boys														
12-17 years	0.64	5.11	1.16	1.01	0.85	0.58	0.78	0.65	0.55	0.67	2.76	1.81	1.99	1.18
12 years	0.66	7.64	1.17	1.07	0.64	1.24	*	---	---	---	---	---	2.11	---
13 years	0.72	*	3.73	1.71	1.89	0.62	0.87	*	---	---	---	---	4.60	---
14 years	0.93	*	*	2.37	1.52	1.32	0.74	1.46	---	*	---	---	3.74	*
15 years	0.74	*	---	*	2.52	2.04	1.25	0.72	1.92	*	*	---	*	*
16 years	0.75	---	---	---	*	3.11	1.38	1.47	0.52	1.61	---	*	*	2.28
17 years	0.79	---	---	---	*	*	1.83	1.89	1.40	0.58	2.96	2.05	*	1.43
Girls														
12-17 years	0.55	4.65	4.09	1.19	0.52	0.66	0.68	0.79	0.46	0.75	2.62	1.74	1.40	0.99
12 years	0.53	6.87	3.51	1.07	0.40	1.13	---	---	---	---	---	---	3.23	---
13 years	0.71	*	*	2.19	1.49	0.52	1.58	*	*	---	---	---	5.25	---
14 years	0.80	*	---	*	2.13	1.54	0.68	1.66	*	---	---	---	*	*
15 years	0.71	---	*	*	*	4.26	1.44	0.84	1.21	*	*	---	*	*
16 years	0.62	---	*	---	*	*	1.44	1.70	0.50	1.83	*	*	*	2.24
17 years	0.75	---	---	---	*	---	*	1.86	1.10	0.68	2.77	1.80	*	1.54
Sample frequencies	6,760	25	74	390	1,056	1,261	1,109	1,072	875	577	55	53	68	145

Table IV. Standard errors for mean raw scores on the Block Design subtest of the WISC by grade, sex, and age for noninstitutionalized youths: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
Both sexes	Standard error													
12-17 years	0.36	2.84	1.72	1.08	0.63	0.40	0.42	0.48	0.52	0.46	1.59	1.61	0.91	1.24
12 years	0.41	3.33	1.58	1.04	0.59	1.48	*	---	---	---	---	---	1.19	---
13 years	0.54	*	4.05	2.28	1.31	0.46	1.09	*	*	---	---	---	2.89	---
14 years	0.56	*	*	3.76	1.74	0.73	0.48	1.31	*	*	---	---	4.00	*
15 years	0.51	*	*	*	2.80	2.41	0.86	0.62	1.25	*	*	---	*	*
16 years	0.56	*	*	---	*	2.78	1.43	1.14	0.67	1.37	*	*	*	2.01
17 years	0.43	---	---	---	*	*	3.03	1.73	1.09	0.52	1.73	1.55	*	1.42
Boys														
12-17 years	0.38	4.99	1.51	1.18	0.76	0.51	0.54	0.68	0.44	0.78	3.11	1.57	1.49	2.52
12 years	0.56	10.16	1.40	1.12	0.71	1.13	*	---	---	---	---	---	1.57	---
13 years	0.62	*	3.26	2.43	1.61	0.63	1.25	*	---	---	---	---	4.35	---
14 years	0.70	*	*	5.07	2.31	1.02	0.71	1.16	---	*	---	---	*	*
15 years	0.72	*	---	*	2.51	2.08	1.08	0.93	2.01	*	*	---	*	*
16 years	0.73	---	---	---	*	3.17	2.19	1.22	0.65	1.54	---	*	*	2.84
17 years	0.51	---	---	---	*	*	3.09	1.55	1.50	0.80	3.17	1.58	*	3.08
Girls														
12-17 years	0.37	1.86	2.85	1.18	0.66	0.51	0.66	0.67	0.77	0.65	3.01	3.25	1.49	1.38
12 years	0.59	2.81	2.98	1.21	0.81	1.97	---	---	---	---	---	---	2.09	---
13 years	0.65	*	*	2.30	1.29	0.67	1.77	*	*	---	---	---	2.83	---
14 years	0.70	*	---	*	2.35	0.95	0.71	2.14	*	---	---	---	*	*
15 years	0.53	---	*	*	*	4.09	1.26	0.80	1.66	*	*	---	*	*
16 years	0.76	*	*	---	*	*	2.03	1.55	1.04	1.74	*	*	*	3.07
17 years	0.56	---	---	---	*	---	*	4.15	1.25	0.81	3.28	3.34	*	2.22
Sample frequencies	6,764	25	74	390	1,057	1,261	1,112	1,071	875	577	55	53	67	147

the range within one standard error of the tabulated statistic with 68 percent confidence, and the range within two standard errors of the tabulated statistic with 95 percent confidence. The latter is used as the level of statistical significance in this report.

An approximation of the standard error of a difference $d = x - y$ of two statistics x and y is given by the formula $S_d = (S_x^2 + S_y^2)^{1/2}$ where S_x and S_y are the sampling errors, respectively, of x and y. Of course, where the two groups or measures are positively or negatively correlated, this formula will give an overestimate or underestimate of the actual standard error.

Thus, the procedure used in this report for testing the significance of difference between means consisted in dividing the difference between the two means by the standard error of the difference as computed above. If the resulting t value was ±2.00 or more, the difference was considered statistically significant at approximately the 5-percent confidence level. For example, the mean Block Design raw score for 12-year-old boys was 25.6, while the mean for 12-year-old girls was 22.4, a difference of 3.2. The approximate standard error of the difference between means was 0.81. Since the difference between the means was about four times the standard error, the difference was considered significant beyond the 5-percent confidence level.

Small Categories

In some tables, statistics may be shown for cells for which the sample size is so small that

Table V. Standard errors for mean scaled scores on the Vocabulary subtest of the WISC by grade, sex, and age for noninstitutionalized youths: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
Both sexes	Standard error													
12-17 years	0.16	1.18	0.62	0.30	0.19	0.14	0.17	0.16	0.13	0.14	0.53	0.33	0.37	0.21
12 years	0.14	1.70	0.59	0.28	0.11	0.28	*	---	---	---	---	---	0.52	---
13 years	0.16	*	1.10	0.52	0.43	0.1	0.26	*	*	---	---	---	0.77	---
14 years	0.22	*	*	0.73	0.35	0.36	0.16	0.31	*	*	---	---	0.72	*
15 years	0.17	*	*	*	0.73	0.68	0.27	0.16	0.35	*	*	---	*	*
16 years	0.16	*	*	---	*	0.49	0.30	0.38	0.09	0.31	*	*	*	0.38
17 years	0.18	---	---	---	*	*	0.60	0.27	0.28	0.14	0.53	0.38	*	0.15
Boys														
12-17 years	0.17	1.47	0.35	0.30	0.24	0.16	0.22	0.18	0.16	0.18	0.86	0.50	0.58	0.30
12 years	0.20	2.35	0.35	0.33	0.19	0.40	*	---	---	---	---	---	0.62	---
13 years	0.19	*	1.03	0.48	0.52	0.16	0.24	*	---	---	---	---	1.37	---
14 years	0.25	*	*	0.52	0.41	0.36	0.20	0.42	---	*	---	---	0.95	*
15 years	0.20	---	---	*	0.68	0.53	0.34	0.19	0.52	*	*	---	*	*
16 years	0.21	---	---	---	*	0.72	0.34	0.41	0.16	0.44	---	*	*	0.59
17 years	0.20	---	---	---	*	*	0.50	0.44	0.36	0.16	0.92	0.55	*	0.32
Girls														
12-17 years	0.15	1.59	1.25	0.35	0.16	0.18	0.18	0.20	0.14	0.21	0.71	0.47	0.37	0.23
12 years	0.15	2.36	1.02	0.31	0.12	0.35	---	---	---	---	---	---	0.91	---
13 years	0.18	*	*	0.61	0.41	0.13	0.42	*	*	---	---	---	1.34	---
14 years	0.21	*	---	*	0.56	0.41	0.19	0.47	*	---	---	---	*	*
15 years	0.18	---	*	*	*	1.20	0.36	0.22	0.30	*	*	---	*	*
16 years	0.17	---	*	---	*	*	0.35	0.48	0.14	0.48	*	*	*	0.52
17 years	0.21	---	---	---	*	---	*	0.50	0.29	0.21	0.75	0.47	*	0.35

the sampling error may be larger than the statistic itself. Such statistics are included in this report, along with their corresponding standard errors, in the belief that the information, while not meeting strict standards of precision, may lend an overall impression of the survey findings and may be of interest to subject-matter specialists.

Quality Control for the Psychological Test Battery

The maintenance of standardized administration and scoring procedures and methods of recording results is essential to large data collection operations such as the Health Examination Survey, as it is to psychological testing itself. Several procedures were incorporated into the administration of the psychological test battery to insure the quality of the data. All psychologists were initially trained in the survey procedures by the Psychological Advisor. In addition, the field psychologists exchanged all test forms daily and checked for any apparent errors in test administration, counting, scoring, or recording. Each field psychologist tape recorded one entire testing session each week. The tapes were sent to the supervisor who reviewed them and made notes of errors and suggestions regarding testing procedures. These notes were then sent to the field psychologists. Periodically the supervisor visited the mobile examination center for direct observation and supervision of the work. The test forms were also spot checked when they arrived at headquarters. Approximately once each week, six WISC subtests, chosen randomly

Table VI. Standard errors for mean scaled scores on the Block Design subtest of the WISC by grade, sex, and age for noninstitutionalized youths: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
	Standard error													
Both sexes														
12-17 years	0.08	1.06	0.48	0.25	0.14	0.09	0.09	0.10	0.12	0.11	0.31	0.41	0.32	0.24
12 years	0.09	1.63	0.42	0.23	0.14	0.33	*	---	---	---	---	---	0.35	---
13 years	0.11	*	1.50	0.51	0.29	0.10	0.24	*	*	---	---	---	0.96	---
14 years	0.12	*	*	0.77	0.34	0.15	0.10	0.26	*	*	---	---	1.08	*
15 years	0.11	*	*	*	0.74	0.52	0.18	0.13	0.26	*	*	---	*	*
16 years	0.11	*	*	---	*	0.52	0.27	0.27	0.15	0.28	*	*	*	0.39
17 years	0.10	---	---	---	*	*	0.59	0.31	0.23	0.12	0.34	0.36	*	0.29
Boys														
12-17 years	0.08	1.40	0.36	0.25	0.17	0.11	0.11	0.16	0.11	0.18	0.75	0.52	0.43	0.51
12 years	0.12	2.76	0.35	0.24	0.17	0.31	*	---	---	---	---	---	0.48	---
13 years	0.13	*	1.44	0.53	0.33	0.13	0.25	*	---	---	---	---	1.33	---
14 years	0.15	*	*	1.02	0.48	0.21	0.16	0.27	---	*	---	---	*	*
15 years	0.15	*	---	*	0.66	0.42	0.21	0.22	0.43	*	*	---	*	*
16 years	0.16	---	---	---	*	0.62	0.41	0.31	0.15	0.33	---	*	*	0.51
17 years	0.11	---	---	---	*	*	0.59	0.30	0.32	0.18	0.76	0.48	*	0.65
Girls														
12-17 years	0.08	1.27	0.90	0.30	0.15	0.11	0.15	0.14	0.17	0.14	0.67	0.71	0.52	0.31
12 years	0.12	1.95	0.83	0.30	0.17	0.41	---	---	---	---	---	---	1.06	---
13 years	0.14	*	*	0.51	0 30	0.14	0.41	*	*	---	---	---	1.71	---
14 years	0.14	*	---	*	0.42	0.21	0.15	0.42	*	---	---	---	*	*
15 years	0.10	---	*	*	*	0.95	0.28	0.16	0.33	*	*	---	*	*
16 years	0.14	*	*	---	*	*	0.34	0.35	0.23	0.36	*	*	*	0.65
17 years	0.13	---	---	---	*	---	*	0.75	0.26	0.17	0.73	0.74	*	0.42

from among those given by each psychologist during the week, were exchanged and rescored. Scoring disagreements were discussed by the two psychologists, and if agreement could not be reached, were referred to the Psychological Advisor for decision. Finally, a log of any unusual occurrences that might affect the validity of scores was maintained.

Table VII. Standard errors for mean scores on the Vocabulary and Block Design short form of the WISC for noninstitutionalized youths, by grade, sex, and age: United States, 1966-70

Sex and age	Total	Present grade in school									High school graduate	Higher than 12th	Special placement	Left before graduating
		Lower than 5th	5th	6th	7th	8th	9th	10th	11th	12th				
Both sexes	Standard error													
12-17 years	0.65	6.00	3.00	1.47	0.91	0.55	0.71	0.66	0.63	0.62	2.17	1.88	1.76	1.17
12 years	0.62	9.26	2.53	1.33	0.62	1.59	*	---	---	---	---	---	1.89	---
13 years	0.76	*	7.27	2.78	2.01	0.48	1.19	*	*	---	---	---	4.51	---
14 years	0.89	*	*	3.94	1.87	1.32	0.68	1.45	*	*	---	---	3.80	*
15 years	0.77	*	*	*	3.13	3.21	1.25	0.79	1.58	*	*	---	*	*
16 years	0.66	*	*	---	*	2.51	1.49	1.60	0.60	1.52	*	*	*	2.12
17 years	0.75	---	---	---	*	*	2.91	1.26	1.36	0.64	2.28	1.92	*	0.95
Boys														
12-17 years	0.70	6.67	1.88	1.46	1.12	0.65	0.80	0.80	0.67	0.87	3.72	2.52	2.67	2.19
12 years	0.83	12.69	1.57	1.42	0.89	1.82	*	---	---	---	---	---	2.74	---
13 years	0.84	*	6.78	2.48	2.33	0.68	0.95	*	---	---	---	---	6.06	---
14 years	1.03	*	*	4.23	2.44	1.45	0.68	1.81	---	*	---	---	16.65	*
15 years	0.95	---	---	*	2.94	2.55	1.45	1.07	2.33	*	*	---	*	*
16 years	0.89	---	---	---	*	3.47	1.97	1.41	0.74	2.06	---	*	*	2.82
17 years	0.81	---	---	---	*	*	2.49	1.71	1.88	0.70	3.83	2.52	*	2.37
Girls														
12-17 years	0.63	8.19	5.78	1.75	0.81	0.74	0.91	0.87	0.75	0.94	3.72	2.93	2.34	1.24
12 years	0.68	12.32	4.46	1.56	0.71	1.85	---	---	---	---	---	---	3.82	---
13 years	0.91	*	*	3.11	1.95	0.62	2.00	*	*	---	---	---	16.55	---
14 years	0.96	*	---	*	2.49	1.53	0.92	*	*	---	---	---	*	*
15 years	0.75	---	*	*	*	5.79	1.74	0.94	1.57	*	*	---	*	*
16 years	0.76	---	*	---	*	*	1.86	2.18	0.90	2.09	*	*	*	3.11
17 years	0.91	---	---	---	*	---	*	3.25	1.47	0.99	4.05	2.92	*	1.99

☆ U. S. GOVERNMENT PRINTING OFFICE : 1973 543-877/9

VITAL AND HEALTH STATISTICS PUBLICATION SERIES

Formerly Public Health Service Publication No. 1000

Series 1. Programs and collection procedures.—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions, data collection methods used, definitions, and other material necessary for understanding the data.

Series 2. Data evaluation and methods research.—Studies of new statistical methodology including: experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, contributions to statistical theory.

Series 3. Analytical studies.—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

Series 4. Documents and committee reports.—Final reports of major committees concerned with vital and health statistics, and documents such as recommended model vital registration laws and revised birth and death certificates.

Series 10. Data from the Health Interview Survey.—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, based on data collected in a continuing national household interview survey.

Series 11. Data from the Health Examination Survey.—Data from direct examination, testing, and measurement of national samples of the civilian, noninstitutional population provide the basis for two types of reports: (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics; and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

Series 12. Data from the Institutional Population Surveys—Statistics relating to the health characteristics of persons in institutions, and their medical, nursing, and personal care received, based on national samples of establishments providing these services and samples of the residents or patients.

Series 13. Data from the Hospital Discharge Survey.—Statistics relating to discharged patients in short-stay hospitals, based on a sample of patient records in a national sample of hospitals.

Series 14. Data on health resources: manpower and facilities.—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

Series 20. Data on mortality.—Various statistics on mortality other than as included in regular annual or monthly reports—special analyses by cause of death, age, and other demographic variables, also geographic and time series analyses.

Series 21. Data on natality, marriage, and divorce.—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports—special analyses by demographic variables, also geographic and time series analyses, studies of fertility.

Series 22. Data from the National Natality and Mortality Surveys.—Statistics on characteristics of births and deaths not available from the vital records, based on sample surveys stemming from these records, including such topics as mortality by socioeconomic class, hospital experience in the last year of life, medical care during pregnancy, health insurance coverage, etc.

For a list of titles of reports published in these series, write to:

Office of Information
National Center for Health Statistics
Public Health Service, HRA
Rockville, Md. 20852

DHEW Publication No. (HRA) 74-1610

Series 11-No. 128

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Public Health Service

HEALTH RESOURCES ADMINISTRATION
5600 Fishers Lane
Rockville, Md. 20852

OFFICIAL BUSINESS
Penalty for Private Use, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF H.E.W.

HEW 396

THIRD CLASS
BLK. RATE

VHSD:BL4M4
MRS SUNDSTROM (2)
ERIC FACILITY
4833 RUGBY AVE
BETHESDA MD 20014